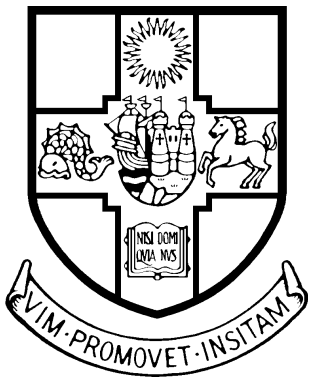
Author:
Paesani, Stefano

Title:
Large-scale integrated quantum photonics
development and applications

LARGE-SCALE INTEGRATED QUANTUM PHOTONICS: DEVELOPMENT AND APPLICATIONS

Stefano Paesani



PhD Thesis

University of Bristol, School of Physics

April 2019

A dissertation submitted to the University of Bristol in accordance with the requirements for award of the degree of Doctor of Philosophy in the Faculty of Sciences.

ABSTRACT

The development of machines that exploit quantum physics to provide fundamentally new computational advantages with respect to conventional or classical hardware is one of most exciting technological prospects for our society. To realise such a computational advantage, these machines will be required to control and process quantum systems at a very large scale, potentially involving millions of quantum information carriers. Single photons, manipulated in linear optical interferometers, are a promising platform to achieve quantum information processing on such a scale. However, the development of photonic technologies able to generate and manipulate photons at this scale has so far proved challenging.

This thesis reports recent progress towards large-scale quantum devices based on integrated quantum photonics. We demonstrate the capability of silicon quantum photonics to integrate photonic circuitry comprising hundreds to thousands of optical components interconnected in stable interferometers. We show how, via the scalable integration of both photon sources and circuitry, such devices can be used to generate, control, and process quantum states of light with an increasing number of photons. Finally, we explore the potential of the developed photonic quantum technology to target key quantum applications, such as the quantum simulation of molecular systems and the learning of Hamiltonian quantum dynamics.

AUTHOR'S DECLARATION

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

SIGNED: ..... DATE: .....

THANKS

Many are the friends and comrades I had the luck to meet during this journey. They have shaped me, and the things I have done. Doing research in Bristol has been a wonderful and exciting experience. I owe it to my supervisors, Anthony and Mark, and to the amazing friends and colleagues I had the privilege to learn from and work with. I am grateful to all members of QETLabs for this. Special thanks are for the companions that were (and, incredibly, still are) patient enough to bear with me in the lab: thanks Raf, Andreas, Massimo, Dan, Seb, Jianwei, Jake and Alex. I am also extremely grateful to Raf, Levon and Jake for going through this thesis and helping me improving it: their views and suggestions are always very precious.

Questi anni sarebbero stati molto piú duri se non stessi facendo ciò che amo fare. Sarò quindi sempre grato a coloro che me lo hanno insegnato. Grazie Prof. Savarese per avermi aperto un universo nuovo. Grazie Lele e Mauro per le discussioni di Fisica tra partitelle di pallone e pranzetti sul prato. Grazie papà, mamma e Giachi: tutto ciò che sono lo devo a voi. Thanks Caterina for sharing this big adventure with me.

Bristol, April 2019 S. P.

LIST OF PUBLICATIONS

The main projects developed during the PhD, which form the main body of the work reported in this thesis, resulted in the following publications

1. Paesani, S., Gentile, A. A., Santagati, R., Wang, J., Wiebe, N., Tew, D. P., O'Brien, J. L. & Thompson, M. G. Experimental Bayesian Quantum Phase Estimation on a Silicon Photonic Chip. *Phys. Rev. Lett.* **118,** 100503. https://link.aps.org/doi/10.1103/PhysRevLett.118.100503 (10 2017).
2. Wang, J., Paesani, S., Santagati, R., Knauer, S., Gentile, A. A., Wiebe, N., Petruzzella, M., O'Brien, J. L., Rarity, J. G., Laing, A., *et al.* Experimental quantum Hamiltonian learning. *Nature Physics* **13,** 551 (2017).
3. Santagati, R., Wang, J., Gentile, A. A., Paesani, S., Wiebe, N., McClean, J. R., Morley-Short, S., Shadbolt, P. J., Bonneau, D., Silverstone, J. W., *et al.* Witnessing eigenstates for quantum simulation of Hamiltonian spectra. *Science Advances* **4,** eaap9646 (2018).
4. Wang, J., Paesani, S., Ding, Y., Santagati, R., Skrzypczyk, P., Salavrakos, A., Tura, J., Augusiak, R., Mančinska, L., Bacco, D., *et al.* Multidimensional quantum entanglement with large-scale integrated optics. *Science* **350,** 285 (2018).
5. Paesani, S., Ding, Y., Santagati, R., Chakhmakhchyan, L., Vigliar, C., Rottwitt, K., Oxenløwe, L. K., Wang, J., Thompson, M. G. & Laing, A. Generation and sampling of quantum states of light in a silicon chip. *Nature Physics* **15,** 925–929 (2019).

During the PhD, the author also appeared in the following additional publications, which are not included in this thesis.

1. Ciampini, M. A., Orieux, A., Paesani, S., Sciarrino, F., Corrielli, G., Crespi, A., Ramponi, R., Osellame, R. & Mataloni, P. Path-polarization hyperentangled and cluster states of photons on a chip. *Light: Science & Applications* **5,** e16064 (2016).
2. Wilkes, C. M., Qiang, X., Wang, J., Santagati, R., Paesani, S., Zhou, X., Miller, D. A., Marshall, G. D., Thompson, M. G. & O'Brien, J. L. 60 dB high-extinction auto-configured Mach–Zehnder interferometer. *Optics Letters* **41,** 5318–5321 (2016).

3. Ciampini, M. A., Vigliar, C., Cimini, V., Paesani, S., Sciarrino, F., Crespi, A., Corrielli, G., Osellame, R., Mataloni, P., Paternostro, M., *et al.* Experimental nonlocality-based network diagnostics of multipartite entangled states. *Scientific Reports* **7,** 17122 (2017).
4. Santagati, R., Gentile, A. A., Knauer, S., Schmitt, S., Paesani, S., Granade, C., Wiebe, N., Osterkamp, C., McGuinness, L. P., Wang, J., Thompson, M. G., Rarity, J. G., Jelezko, F. & Laing, A. Magnetic-Field Learning Using a Single Electronic Spin in Diamond with One-Photon Readout at Room Temperature. *Phys. Rev. X* **9,** 021019. https://link.aps.org/doi/10.1103/PhysRevX.9.021019 (2 2019).
5. Llewellyn, D., Ding, Y., Faruque, I. I., Paesani, S., Bacco, D., Santagati, R., Qian, Y.-J., Li, Y., Yun-Feng, X., Huber, M., Malik, M., Sinclair, G. F., Zhou, X., Rottwitt, K., O'Brien, J. L., Rarity, J. G., Gong, Q., Oxenløwe, L. K., Wang, J. & G., T. M. Optical Microresonator-enhanced Multiqubit Quantum Information Processing. *(Submitted)* (2019).

The work performed is also reported in the following conference proceedings.

1. Santagati, R., Wang, J., Paesani, S., Gentile, A. A., Bonneau, D., Silverstone, J., Morley-Short, S., Zhou, X., Shadbolt, P., Wiebe, N., *et al. Ground state solver on a silicon quantum photonic chip* in *CLEO: QELS_Fundamental Science* (2016), FTh4C–3.
2. Paesani, S., Gentile, A. A., Santagati, R., Wang, J., Wiebe, N., Tew, D., O'Brien, J. L. & Thompson, M. G. *Bayesian quantum phase estimation on an integrated Silicon photonic device* in *Frontiers in Optics* (2016), FTu3G–5.
3. Wilkes, C. M., Qiang, X., Wang, J., Santagati, R., Paesani, S., Zhou, X., Shadbolt, P., Rudolph, T., Miller, D., Thompson, M. G., *et al. 60.5 dB Silicon Mach-Zehnder Interferometer using Self-Optimising Beam-Splitters* in *Frontiers in Optics* (2016), FTu1D–3.
4. Wang, J., Paesani, S., Santagati, R., Knauer, S., Gentile, A. A., Wiebe, N., Petruzzella, M, Laing, A., Rarity, J. G., O'Brien, J. L., *et al. Learning nitrogen-vacancy electron spin dynamics on a silicon quantum photonic simulator* in *2017 Conference on Lasers and Electro-Optics (CLEO)* (2017), 1–1.
5. Gentile, A., Paesani, S, Santagati, R, Wang, J, Wiebe, N, Tew, D, O'Brien, J. & Thompson, M. *Noise resilience of Bayesian quantum phase estimation tested on a Si quantum photonic chip* in *European Quantum Electronics Conference* (2017), EB_6_2.
6. Santagati, R., Wang, J., Gentile, A., Paesani, S., Wiebe, N., McClean, J., Bonneau, D., Silverstone, J., Morley-Short, S., Shadbolt, P., *et al. Finding excited states of physical Hamiltonians on a silicon quantum photonic device* in *Frontiers in Optics* (2017), FM4E–2.
7. Santagati, R., Wang, J., Paesani, S., Knauer, S., Gentile, A., Wiebe, N., Petruzzella, M., O'Brien, J., Rarity, J., Laing, A., *et al. Towards practical characterization of quantum systems with quantum Hamiltonian learning* in *Frontiers in Optics* (2017), FTh3E–7.

8. Paesani, S., Wang, J., Santagati, R., Knauer, S., Gentile, A. A., Wiebe, N., Petruzzella, M., Laing, A., Rarity, J. G., O'Brien, J. L., *et al. Experimental quantum hamiltonian learning using a silicon photonic chip and a nitrogen-vacancy electron spin in diamond* in *Lasers and Electro-Optics Europe & European Quantum Electronics Conference (CLEO/Europe-EQEC, 2017 Conference on)* (2017), 1–1.
9. Ciampini, M., Orieux, A, Paesani, S, Vigliar, C, Cimini, V, Corrielli, G, Crespi, A, Ramponi, R, Osellame, R, Paternostro, M, *et al. On chip analysis of path-polarization hyperentangled cluster photon states* in *Advances in Photonics of Quantum Computing, Memory, and Communication X* **10118** (2017), 101180I.
10. Wang, J., Paesani, S., Ding, Y., Santagati, R., Skrzypczyk, P., Salavrakos, A., Tura, J., Augusiak, R., Mancinska, L., Bacco, D., *et al. Large-scale Integration of Multidimensional Quantum Photonics Circuits on Silicon* in *CLEO: Science and Innovations* (2018), JTh5B–4.
11. Faruque, I., Llewellyn, D, Ding, Y, Paesani, S, Santagati, R, Bonneau, D, Sinclair, G., Bacco, D., Rottwitt, K., Oxenløwe, L. K., *et al. Indistinguishable Photon-Pairs from Pure and Bright Silicon Micro-ring Resonator Sources* in *CLEO: QELS_Fundamental Science* (2018), FTh1G–5.
12. Llewellyn, D, Ding, Y, Faruque, I, Paesani, S, Santagati, R, Kennard, J, Bacco, D., Rottwitt, K., Oxenløwe, L. K., O'Brien, J., *et al. Generation and Manipulation of Multi-Photon Entangled States on a Silicon Photonic Device* in *CLEO: QELS_Fundamental Science* (2018), FTh1G–7.

CONTENTS

1 INTRODUCTION

The development of quantum mechanics has been one of the greatest scientific achievements of the 20^{th} century. Up to now, it represents the most complete and accurate theory to describe the physical world, and over the past decades it has become increasingly evident that it strongly underpins a wide variety of scientific fields, from biology to astrophysics. Probably, the most intriguing peculiarity of quantum mechanics is its introduction of concepts that are highly counter-intuitive. It requires us to accept rules that are completely new with respect to our standard classical description of nature.

Scientists have realised that by taking advantage of the counter-intuitive features of quantum mechanics, technologies with fundamentally new capabilities are possible. Quantum mechanics has already driven a first technological revolution with the development of, for example, the transistor, light-emitting diodes (LEDs), and lasers. These are all technologies based on quantum effects arising in the collective behaviour of weakly controlled quantum systems.

However, scientific advancements have now brought us to a point where individual quantum entities can be generated and manipulated in a very controlled environment, giving access to the most fundamental properties of quantum systems, such as quantum entanglement. Machines based on these effects promise to drive a second quantum technological revolution. Many new applications are emerging for such machines: quantum computing, quantum simulation, quantum sensors, and quantum cryptography are the main examples. However, the challenge of engineering, controlling and validating quantum devices is significant, and the pursuit for practical quantum advantages represents one of the greatest current scientific efforts. A wide variety of different approaches and platforms are being investigated, which includes ion trap systems, quantum dots, superconducting circuits, nitrogen-vacancy centres in diamonds, and many others.

This thesis is a contribution to the development of quantum technologies based on photonics: the control of single particles of light. Quantum photonics has already played,

and is still playing, a great role in testing and exploring fundamental properties of quantum mechanics. However, when aiming at developing quantum technologies based on photons, significant new challenges arise. For example, the generation, control and processing of the number of photons required in practical quantum computing applications demands the implementation of millions of optical components, including photon sources, interferometers, and detectors; a task which is hardly conceivable with standard optical approaches. Fortunately, integrated optics provides a more viable route. In particular, in this thesis we aim at using the silicon quantum photonics platform for the development of large-scale photonic quantum technologies. During the thesis we will report experimental progress in this direction, and discuss and demonstrate quantum applications that can be addressed with such devices.

1.1 Thesis outline

The structure of this thesis can be divided into three parts. The first, which includes chapter 2, is introductory to the basic concepts and notations that will be used in the research parts of the thesis. The second part (chapters 3 and 4) is focused on the technological progress of the silicon quantum photonics platform. In the last part (chapters 5, 6, and 7) we will discuss possible applications for the quantum photonics hardware developed, with a main focus on quantum simulation of chemical systems. The content of the different chapters is as follows.

- **Background part.** Chapter 2 provides a general description of the platform used: silicon quantum photonics. After an initial introduction on the description of quantum states of light and quantum optical processes, we will discuss the architectures and circuits required for photonic quantum information processing. The integrated photonics platform will be discussed, and the main integrated optical components used to generate and manipulate photons on silicon devices will be introduced.

- **Technology part.** Chapter 3 concerns progress towards the large-scale integration of optical circuits in a single photonic integrated device. We present an experiment involving a silicon photonic chip which embeds probably the most complex quantum photonic circuit reported to date. Such circuit is used to generate and process high-dimensional states of entangled photons on a chip. Chapter 4 is instead focused on increasing the number of photons, the main source of computational complexity in quantum photonics. We will report an experiments where up to eight photons are generated and processed on a silicon chip, currently a record in integrated quantum optics. These photons will be used to demonstrate different

boson sampling protocols.

- **Applications part.** In chapter 5 we discuss the analog quantum simulation of molecular processes using photons. In particular, we report experimental results where we benchmark our silicon photonics technology for the simulation of molecular Franck-Condon profiles via boson sampling. We proceed in Chapter 6 to digital quantum simulation using photons. We first overview quantum algorithms for the efficient calculation of energies in quantum chemistry, and develop a new quantum variational algorithm for determining the energies associated to molecular excited states. Small-scale integrated quantum photonics implementations of the discussed quantum algorithms are reported. Finally, in 7 we will combine quantum simulation with machine learning ideas, and discuss a protocol which allows us to efficiently learn quantum Hamiltonians using quantum simulation. An experimental demonstration of the protocol is reported, where an integrated photonic quantum simulator is used to learn the Hamiltonian of an electron spin in a colour centre in diamond.

2 QUANTUM INFORMATION PROCESSING IN INTEGRATED LINEAR OPTICS

Starting from Euclid's Optics in 300 BC, light has always played a central role in human's understanding of Nature, allowing us to probe the most fundamental properties of physical objects and driving countless scientific discoveries. Quantum mechanics itself originates from the study of the black body radiation and the photo-electric effect. Light is also playing an important role in the development of quantum technologies.

Using photons as quantum information carriers represents a completely different paradigm compared to other quantum architectures, bringing both advantages and drawbacks. Firstly, the platform is unique in being based on *flying qubits*, i.e. qubits capable of travelling long distances. This contrasts with all other approaches, e.g., on solid-state or trapped-ions, where qubits are localised in space. Such property is extremely useful for connecting different modules of a quantum computer, and in technologies such as quantum communication. Remarkably, photons also constitute the only quantum system which does not decohere. This property is manifested in the fact that the cosmic background radiation, originated in the early stages of the Big Bang, still possess a coherent polarisation [1], indicating that the decoherence time of photons is *at least* as long as the age of the Universe. The absence of decoherence is due to the difficulty in having photons interact. On the other hand, this feature precludes to directly implement gates between them. As we will see in this chapter, this issue requires clever schemes where the photon-photon interactions are mediated via the action of measurements, granting the possibility of scalable universal schemes for linear-optics quantum computing. Moreover, photon loss is a type of error which uniquely arises in quantum photonics, and represents the main type of noise in this platform. The error models and the computational architectures required for quantum photonics are thus quite different from what is typically addressed in the quantum computing community.

This chapter is intended to provide a general background on photonic quantum information processing. We will start by outlining the standard notation to represent quantum states of light, and describe how they can be generated and processed with linear optical components. We will then pass to discuss a promising technology to target near-term quantum applications with photons: integrated optics. This technology will be used in the rest of the thesis to implement large-scale quantum photonics experiments.

2.1 Quantum states of light and their evolution

In the second quantization formalism, light is described as an ensemble of quantum harmonic oscillators with Hamiltonian

$$\hat{H} = \sum_k \hat{H}_k, \qquad \hat{H}_k = \hbar\omega_k \left(\hat{a}_k^\dagger \hat{a}_k + \frac{1}{2}\right), \tag{2.1.1}$$

where each Hamiltonian $\hat{H}_k$ refers to a single harmonic oscillator with angular frequency ω_k associated to the k-th mode of the electromagnetic field. These modes can be, for example, spatial modes, spectral modes, polarisations, and so on. In this thesis, unless specified otherwise, we will generally consider spatial modes: different waveguides in a photonic chip, for example. The operators $\hat{a}_k^\dagger$ and $\hat{a}_k$ are bosonic creation and annihilation operators for excitations in mode k, and obey the bosonic commutation rules

$$\left[\hat{a}_k, \hat{a}_\ell\right] = \left[\hat{a}_k^\dagger, \hat{a}_\ell^\dagger\right] = 0, \qquad \left[\hat{a}_k^\dagger, \hat{a}_\ell\right] = -\delta_{k\ell}. \tag{2.1.2}$$

The analog of the position and momentum operators for the k-th mode are the *quadrature operators* $\hat{q}_k$ and $\hat{p}_k$, which are related to the creation and annihilation operators via

$$\hat{q}_k = \left(\hat{a}_k + \hat{a}_k^\dagger\right)/\sqrt{2}, \qquad \hat{p}_k = \left(\hat{a}_k - \hat{a}_k^\dagger\right)/i\sqrt{2} \tag{2.1.3}$$

and satisfy $[\hat{q}_k, \hat{p}_\ell] = i\delta_{k\ell}$. These space-like and momentum-like operators can be used to describe the states of light in the *phase space*. This space has a symplectic structure described by the commutation rule

$$[\hat{x}_k, \hat{x}_\ell] = i\mathbf{\Omega}_{k\ell}, \tag{2.1.4}$$

where, if the system has m total modes, $\hat{\boldsymbol{x}} = (\hat{q}_1, \ldots, \hat{q}_m, \hat{p}_1, \ldots, \hat{p}_m)$ is the vector of the quadrature operators and $\mathbf{\Omega}_{k\ell}$ are the elements is the $2m \times 2m$ matrix

$$\mathbf{\Omega} = \begin{bmatrix} 0 & \mathbb{1}_m \\ -\mathbb{1}_m & 0 \end{bmatrix}. \tag{2.1.5}$$

A matrix $\boldsymbol{S}$ acting on this space is called *symplectic* if it satisfies $\boldsymbol{S}\boldsymbol{\Omega}\boldsymbol{S}^\top = \boldsymbol{\Omega}$. The vector $\hat{\boldsymbol{x}} = (\hat{q}_1, \dots, \hat{q}_m, \hat{p}_1, \dots, \hat{p}_m)$ forms an operator basis, called *quadrature basis*. A different basis which is often convenient is the complex *Heisenberg basis* $\hat{\boldsymbol{\xi}}^{(H)} = (\hat{a}_1, \dots, \hat{a}_m, \hat{a}_1^\dagger, \dots, \hat{a}_m^\dagger)$. The representation $\boldsymbol{S}^{(H)}$ of a symplectic matrix in this basis is given by

$$\boldsymbol{S}^{(H)} = \boldsymbol{\Lambda}_0 \boldsymbol{S} \boldsymbol{\Lambda}_0^\dagger, \qquad \text{with} \qquad \boldsymbol{\Lambda}_0 = \frac{1}{\sqrt{2}} \begin{bmatrix} \mathbb{1}_m & i\mathbb{1}_m \\ \mathbb{1}_m & -i\mathbb{1}_m \end{bmatrix}. \tag{2.1.6}$$

2.1.1 *Single photons*

Photons represent the single bosonic quantum excitations of the electromagnetic field. The photonic states are indicated via Fock vectors[1] $|n\rangle_k$, where n represents the number of photons on mode k, and satisfy

$$\hat{a}_k^\dagger |n\rangle_k = \sqrt{n+1}\,|n+1\rangle\,, \qquad \hat{a}_k |n\rangle_k = \sqrt{n}\,|n-1\rangle\,, \qquad \hat{N}_k = \hat{a}_k^\dagger \hat{a}_k, \tag{2.1.7}$$

where $\hat{N}_k$ is the *number operator* on mode k. Photons occupying the same optical mode are said to be *indistinguishable*, and are *distinguishable* otherwise. If mode k is occupied by n photons, the energy of the electromagnetic field associated to it is $\langle n|_k \hat{H} |n\rangle_k = \hbar\omega_k(n + 1/2)$. The state $|0\rangle_k$, where no photons are present in the k-th mode, is the *vacuum state* associated to mode k. Any Fock state can be written as

$$|n\rangle_k = \frac{1}{\sqrt{n!}} \left(\hat{a}_k^\dagger\right)^n |0\rangle_k\,. \tag{2.1.8}$$

In a configuration where a total of $n = \sum_{k=1}^m n_k$ photons are in m different modes, where the k-th mode contains n_k photons, the total Fock state of the system is given by

$$|\boldsymbol{n}\rangle = |n_1 n_2 \dots n_m\rangle = |n_1\rangle_1 \otimes |n_2\rangle_2 \otimes \dots \otimes |n_m\rangle_m\,, \tag{2.1.9}$$

where we have defined $\boldsymbol{n} = (n_1, n_2, \dots, n_m)$. This class of states form an orthogonal basis of the Hilbert space $\mathcal{H}_m^n$ of n indistinguishable photons in m modes, which has dimension

$$\dim\left(\mathcal{H}_m^n\right) = \binom{m+n-1}{n} \tag{2.1.10}$$

1 Fock states are indicated using the same bra-ket notation as for logical states of, e.g., qubits. The meaning of the two is however different in general, and should not be confused. In this thesis the meaning of the notation adopted should be clear from the context in which it is used, or specified otherwise.

equal to the number of possible configurations for allocating n indistinguishable objects in m slots. An arbitrary state of n photons in m modes can thus be written as the superposition

$$|\psi\rangle = \sum_{\boldsymbol{n}} \beta_{\boldsymbol{n}} |\boldsymbol{n}\rangle , \tag{2.1.11}$$

where $\boldsymbol{n}$ ranges over all possible configurations $(n_1, n_2, \ldots, n_m)$ with fixed total number of photons $n = \sum_k n_k$, and $\sum_k |\beta_k|^2 = 1$.

2.1.2 *Coherent states*

A second type of quantum state of light are *coherent states* $|\alpha\rangle$, These are defined as the right-eigestates of the annihilation operator $\hat{a}_k |\alpha\rangle_k = \alpha |\alpha\rangle_k$, where α is a complex number representing the *displacement*. Coherent states can be obtained by applying the displacement operator $\hat{D}_k(\alpha)$ to the vacuum: $|\alpha\rangle_k = \hat{D}_k(\alpha) |0\rangle_k$, where

$$\hat{D}_k(\alpha) = \exp\Big(\alpha \hat{a}_k^\dagger - \alpha^* \hat{a}_k\Big). \tag{2.1.12}$$

The action of the displacement on the mode operators is given by $\hat{a}_k \mapsto \hat{a}_k + \alpha_k$. Multi-mode displacement can be compactly written as the operator

$$\hat{D}(\boldsymbol{\lambda}) = \exp\Big(\hat{\boldsymbol{x}}^\top \boldsymbol{\Omega} \boldsymbol{\lambda}\Big), \tag{2.1.13}$$

where $\boldsymbol{\lambda} = \sqrt{2}(\operatorname{Re}\alpha_1, \operatorname{Im}\alpha_1, \ldots, \operatorname{Re}\alpha_m, \operatorname{Im}\alpha_m)$. Coherent states admit a simple and direct analytical decomposition in the Fock basis:

$$|\alpha\rangle_k = e^{-\frac{|\alpha|^2}{2}} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle_k . \tag{2.1.14}$$

The probability of detecting n photons from a coherent state is then given by

$$p(n|\alpha) = |\langle n|\alpha\rangle|^2 = e^{-|\alpha|^2} \frac{|\alpha|^{2n}}{n!}, \tag{2.1.15}$$

which is a Poissonian distribution with average photon number $\langle n\rangle = |\alpha|^2$.

Coherent states can be used to describe, with a good degree of approximation, the emitted light from a laser. Broadly speaking, they are the quantum states of light with higher resemblance to the dynamics of a classical harmonic oscillator, and are therefore often regarded as classical [2].

2.1.3 *Linear optics*

Most of the dynamics of the electromagnetic field, e.g. in absence of light-matter interactions, can be described as a linear transformation of the mode operators [2–4]:

$$\hat{a}_k^\dagger \mapsto \sum_\ell c_{k\ell}^{(1)} \hat{a}_\ell^\dagger + c_{k\ell}^{(2)} \hat{a}_\ell + c_k^{(3)}, \tag{2.1.16}$$

also known as the *Bogoliubov transformation*. Such evolution can in general be described using an Hamiltonian of the form:

$$\hat{H} = \underbrace{\sum_{k=1}^{m} \alpha_k \hat{a}_k^\dagger + h.c.}_{\text{(1) Displacement}} + \underbrace{\sum_{k,\ell=1}^{m} \xi_{k\ell} \hat{a}_k^\dagger \hat{a}_\ell^\dagger + h.c.}_{\text{(2) Squeezing}} + \underbrace{\sum_{k\geq\ell=1}^{m} \mu_{k\ell} \hat{a}_k^\dagger \hat{a}_\ell + h.c.}_{\text{(3) Linear optics}} \tag{2.1.17}$$

Three terms can be distinguished in this Hamiltonian, each representing a different building block of quantum optics. Note that the first two terms, (1) and (2), do not conserve the total number of photons. In particular, the term (1) has the form of the displacement operator (2.1.12), and therefore represents an evolution which displaces the electromagnetic field in the associated mode. The term (2) is instead associated to *squeezing*, as we will shortly see.

In this thesis we will mainly focus on the third term (3), which represents *linear optical* transformations: fully passive evolutions that maintain the total number of photons. Denoting as $\mathcal{U} = \exp(-iH_3 t)$ the unitary map associated to the term (3), the transformation on the mode operators $\hat{\boldsymbol{a}}^\dagger = (\hat{a}_1^\dagger, \dots, \hat{a}_m^\dagger)$ can be formulated, in the Heisenberg picture, as a simple matrix multiplication:

$$\hat{\boldsymbol{a}}^\dagger \mapsto \mathcal{U} \hat{\boldsymbol{a}}^\dagger \mathcal{U}^\dagger = U^\top \hat{\boldsymbol{a}}^\dagger, \tag{2.1.18}$$

where U is a $m \times m$ unitary matrix given by $U = \exp(-i\mu t)$ and μ is the Hermitian matrix with elements $\mu_{k\ell}$ from eq.(2.1.17).

Evolutions in a linear optical system with m modes are thus described via $m \times m$ unitary matrices. The unitarity of the matrix has the meaning of conserving the total energy in the system. An interpretation of the matrix elements of U can be immediately derived from the following observation. If a photon is prepared in the i-th mode and evolves through a linear optical transformation $\mathcal{U}$, the probability of detecting it in the j-th output mode is given by

$$p(j|i) = |\langle \mathbf{0}| \hat{a}_j \left(\mathcal{U} \hat{a}_i^\dagger \mathcal{U}^\dagger\right) |\mathbf{0}\rangle|^2 = |\langle \mathbf{0}| \hat{a}_j \sum_k U_{ki} \hat{a}_k^\dagger |\mathbf{0}\rangle|^2 = |U_{ji}|^2. \tag{2.1.19}$$

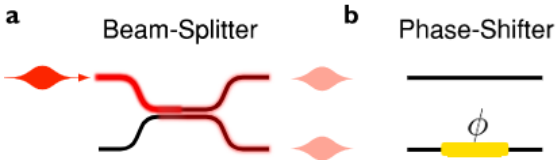


Figure 2.1: Building blocks for linear optics. **a,** The beam-splitter is used to mix the amplitudes of different modes. **b,** The phase shifter inserts an optical phase ϕ in the associated mode.

The squared module on the element of the j-th row and i-th column of U can thus be considered as the transition probability for a single photon from mode i to j.

The action of a linear optical element on a coherent state can be described similarly. In fact it is easy to see that a linear optical transformation maps any coherent states into another coherent state with amplitudes $\mathcal{U}|\boldsymbol{\alpha}\rangle = |U\boldsymbol{\alpha}\rangle$. In particular, if coherent light is injected in mode i, the amplitude in the j-th output mode is given by a rescaling factor $|U_{ji}|^2$ analog to the single photon case. An individual photon thus has the same behaviour in a linear optical network as a classical coherent state. As we will see in the next section, the difference arises when more than one photon is used, due to quantum interference.

Linear optical operations can be constructed using two building blocks. The first is an intra-mode operation, i.e. with $k = \ell$ in the term (3) of eq.(2.1.17), which represents *phase shifters*. The second is a inter-mode operation, i.e. $k \neq \ell$, representing *beam-splitters*. In a two-mode scenario (see Fig. 2.1), their corresponding unitary evolutions are respectively given by

$$U_{\mathrm{PS}}(\phi) = \begin{bmatrix} 1 & 0 \\ 0 & e^{i\phi} \end{bmatrix}, \qquad U_{\mathrm{BS}}(\eta) = \begin{bmatrix} \sqrt{\eta} & i\sqrt{1-\eta} \\ i\sqrt{1-\eta} & \sqrt{\eta} \end{bmatrix}, \tag{2.1.20}$$

where ϕ is the phase applied (on the second mode) by the phase shifter, and η is the the beam-splitter reflectivity. In the next section we will describe how any unitary linear optical operation can be performed on an arbitrary number of modes by cascading these two building blocks.

2.1.4 *Gaussian states of light*

In addition to single photons, another broad class of quantum states of light which plays an important role in quantum optics and quantum information are *Gaussian states*. They will be extensively used in chapters 4 and 5. Although their definition may seem a bit abstract and intricate at first sight, we will see how the formalism developed allows us to describe Gaussian states and their evolution in an elegant and quite simple manner.

Gaussian states are essentially states of light that allow a simple representation in phase space via multivariate Gaussian distributions. More formally, to describe a general quantum state ρ in phase space, we can use the *characteristic function*

$$\chi(\boldsymbol{\lambda}) = \mathrm{Tr}\left[\rho \hat{D}(\boldsymbol{\lambda})\right], \tag{2.1.21}$$

which essentially gives the expectation value of arbitrary coherent states (see eq. 2.1.13) over ρ and is related to the Wigner function via a Fourier transform [2]

$$\mathcal{W}(\boldsymbol{x}) = \frac{1}{(2\pi)^{2m}} \int_{\mathbb{R}^{2m}} d\boldsymbol{\lambda}\ \exp\Big(-i\boldsymbol{x}^\top \boldsymbol{\Omega}\boldsymbol{\lambda}\Big)\chi(\boldsymbol{\lambda}). \tag{2.1.22}$$

The Wigner function is a quasi-probability distribution (sums to one but can be negative) describing the state in phase space. Its first two moments are the displacement $\bar{\boldsymbol{x}}$ and the *covariance matrix* $\boldsymbol{\sigma}$, given by:

$$\bar{\boldsymbol{x}} = \langle \hat{\boldsymbol{x}} \rangle = \mathrm{Tr}\left(\rho \hat{\boldsymbol{x}}\right), \qquad \sigma_{ij} = [\boldsymbol{\sigma}]_{ij} = \frac{1}{2}\langle \hat{x}_i \hat{x}_j + \hat{x}_j \hat{x}_i \rangle - \bar{x}_i \bar{x}_j. \tag{2.1.23}$$

Gaussian states are defined as the states with characteristic and Wigner function given by multivariate Gaussian distributions in the phase space, and can thus be written as [2, 4, 5]

$$\chi(\boldsymbol{\lambda}) = \exp\left[-\frac{1}{2}\boldsymbol{\lambda}^\top \boldsymbol{\Omega}\boldsymbol{\sigma}\boldsymbol{\Omega}^\top \boldsymbol{\lambda} - i\boldsymbol{\lambda}^\top \boldsymbol{\Omega}\bar{\boldsymbol{x}}\right], \tag{2.1.24}$$

$$\mathcal{W}(\boldsymbol{x}) = \frac{1}{\pi^m \sqrt{\det \boldsymbol{\sigma}}} \exp\left[-\frac{1}{2}\left(\boldsymbol{x} - \bar{\boldsymbol{x}}\right)^\top \boldsymbol{\sigma}^{-\mathbf{1}} \left(\boldsymbol{x} - \bar{\boldsymbol{x}}\right)\right]. \tag{2.1.25}$$

Gaussian states are thus fully characterised by the first two moments. Note that, for all Gaussian states, the Wigner function is necessarily always positive.

For convenience, in this thesis we will always represent the covariance matrices using the Heisenberg basis (we will drop the (H) label for simplicity), which are obtained from the definition of eq.(2.1.23) by applying the transformation in eq.(2.1.6). In the Heisenberg basis, the covariance matrix is Hermitian, positive-definite, and must satisfy the uncertainty relation $\boldsymbol{\sigma} + \frac{1}{2}\boldsymbol{Z}_{2m} \geq 0$, where

$$\boldsymbol{Z}_{2m} = \begin{bmatrix} \mathbb{1}_m & 0 \\ 0 & -\mathbb{1}_m \end{bmatrix} \tag{2.1.26}$$

and ≥ 0 indicates semi-positiveness. Pure Gaussian states are the ones with minimum variance, i.e. that saturate the uncertainty bound $\boldsymbol{\sigma} + \frac{1}{2}\boldsymbol{Z}_{2m} = 0$ [2].

Apart from being easy to characterise, Gaussian states also allow us to describe their transformations in a very simple way. A transformation is said to be Gaussian when it maps Gaussian states into Gaussian states. Gaussian transformations are characterised by linear operations on the moments [2]:

$$\bar{\boldsymbol{x}} \mapsto \boldsymbol{S}\bar{\boldsymbol{x}} + \boldsymbol{d}, \qquad \boldsymbol{\sigma} \mapsto \boldsymbol{S}\boldsymbol{\sigma}\boldsymbol{S}^\dagger, \tag{2.1.27}$$

where $\boldsymbol{S}$ is a symplectic matrix associated to the transformation. The importance of these states in quantum optics comes from the fact that a large part of optical operations are Gaussian. In particular, operations are Gaussian *if and only if* they are generated from Hamiltonians as eq.(2.1.17), which encompasses most transformations in quantum optics. The symplectic matrices associated to all terms in eq.(2.1.17) can be readily obtained, as described below.

Clearly, not all states and transformations in quantum optics are Gaussians. Simple examples are single photon states and single photon measurements. However, a large class of elements in quantum optics allow a simple Gaussian representation. Below some of the basic Gaussian states, which will be used in chapter 4 and 5, are described. For a more formal derivation of their properties the reader can refer to, for example, reference [2].

Vacuum and thermal states

In the Gaussian formalism, the representation in phase space of the vacuum state is simply given by

$$\bar{\boldsymbol{x}}_{\text{vac}} = \boldsymbol{0}, \qquad \boldsymbol{\sigma}_{\text{vac}} = \mathbb{1}_{2m}/2. \tag{2.1.28}$$

All Gaussian states with zero-displacement arise from vacuum. A displacement $\bar{\boldsymbol{x}} = \bar{\boldsymbol{d}}$ can be added via the displacement transformation, which, in the Gaussian formalism of eq.(2.1.27), is given by

$$\boldsymbol{d}_{\text{coh}} = \bar{\boldsymbol{d}}, \qquad \boldsymbol{S}_{\text{coh}} = \mathbb{1}_{2m}. \tag{2.1.29}$$

The covariance matrix associated to a coherent state in the Gaussian formalism is then simply given by $\boldsymbol{\sigma}_{\text{coh}} = \boldsymbol{S}_{\text{coh}}\boldsymbol{\sigma}_{\text{vac}}\boldsymbol{S}_{\text{coh}}^{\dagger} = \mathbb{1}_{2m}/2$.

The vacuum and coherent states are pure Gaussian states, i.e. with minimal variance [2]. An important class of mixed Gaussian states are *thermal states*, that are the states characterising, for example, the black-body radiation. In the photon number basis, a thermal state in a single mode is given by the mixture

$$\rho_{\text{thm}} = \frac{1}{\bar{n}+1}\sum_{n=0}^{\infty}\left(\frac{\bar{n}}{\bar{n}+1}\right)^{n}|n\rangle\langle n|, \tag{2.1.30}$$

with $\bar{n} = \langle n\rangle$ the average number of photons. The photon number statistics in a thermal state are distributed according to the Bose-Einstein distribution describing black-body radiation:

$$p_{\text{thm}}(n) = \langle n|\rho_{\text{thm}}|n\rangle = \frac{1}{\bar{n}+1}\left(\frac{\bar{n}}{\bar{n}+1}\right)^{n} = e^{-\frac{\hbar n\omega}{k_B T}}\left(1-e^{-\frac{\hbar\omega}{k_B T}}\right), \tag{2.1.31}$$

where T is the temperature of the associated black-body emitting the radiation, k_B the Boltzmann constant. For the Bose-Einstein distribution the average number of photons

is given by $\bar{n} = (e^{\hbar\omega/k_B T} - 1)^{-1}$. Vacuum states correspond to thermal states at zero temperature. It is easy to check that thermal states are Gaussian states, with the first and second moments given by:

$$\bar{\boldsymbol{x}}_{\text{thm}} = \mathbf{0}, \qquad \boldsymbol{\sigma}_{\text{thm}} = \frac{1}{2}\left[\text{diag}(2\bar{\boldsymbol{n}} + \mathbf{1}) \oplus \text{diag}(2\bar{\boldsymbol{n}} + \mathbf{1})\right]. \tag{2.1.32}$$

Linear optical transformations in the Gaussian formalism

The action of linear optical transformation on the mode operators is quite simply given by $\hat{\boldsymbol{a}}^\dagger \mapsto U^\top \hat{\boldsymbol{a}}^\dagger$, with U a unitary matrix. As a consequence, the symplectic operation associated to linear optical transformations in the Heisenberg basis is therefore of the form

$$S_{\text{LO}}(U) = \begin{bmatrix} U & 0 \\ 0 & U^* \end{bmatrix}, \tag{2.1.33}$$

If an input Gaussian state with covariance matrix $\boldsymbol{\sigma}$ is injected into a linear optical network, the output state will then be simply described by

$$\boldsymbol{\sigma} \mapsto \begin{bmatrix} U & 0 \\ 0 & U^* \end{bmatrix} \boldsymbol{\sigma} \begin{bmatrix} U^\dagger & 0 \\ 0 & U^\top \end{bmatrix}. \tag{2.1.34}$$

Single-mode squeezing

So far we have discussed the terms (1) and (3), corresponding to displacement and linear optics, respectively. Having discussed the formalism of Gaussian states, it is now simple to characterise the term (2), which corresponds to squeezing. We can start considering the *single-mode squeezing* (SMS) case, which corresponds to the cases with same indices ($k = \ell$) in eq.(2.1.17). The SMS operator is of the form

$$\mathcal{S}_{\text{SMS}}(\xi) = \exp\left[\frac{1}{2}\left(\xi(\hat{a}^\dagger)^2 - \xi^* \hat{a}^2\right)\right], \tag{2.1.35}$$

where $\xi = re^{i\varphi}$ is the *squeezing parameter*. The mode evolutions corresponding to this operator are given by

$$\hat{a} \mapsto \hat{a}\cosh r + \hat{a}^\dagger e^{i\varphi}\sinh r, \tag{2.1.36}$$

which, in the case where SMS is performed on m modes, induces the symplectic matrix

$$S_{\text{SMS}} = \begin{bmatrix} \bigoplus_{k=1}^m \cosh r_k & \bigoplus_{k=1}^m e^{i\varphi_k}\sinh r_k \\ \bigoplus_{k=1}^m e^{-i\varphi_k}\sinh r_k & \bigoplus_{k=1}^m \cosh r_k \end{bmatrix}. \tag{2.1.37}$$

The states generated by the single-mode squeezing symplectic matrix acting on the vacuum are called *vacuum single-mode squeezed states*, and have zero displacement and covariance matrix $\boldsymbol{\sigma}_{\mathrm{SMS}} = \frac{1}{2} S_{\mathrm{SMS}} S_{\mathrm{SMS}}^{\dagger}$. Single-mode squeezed states are pure, and can be written in terms of single photon states as

$$|\mathrm{SMS}(\xi_k)\rangle_k = \frac{1}{\sqrt{\cosh r_k}} \sum_{n=0}^{\infty} (-e^{i\varphi_k} \tanh r_k)^n \frac{\sqrt{(2n)!}}{2^n n!} |2n\rangle_k . \tag{2.1.38}$$

Note that only terms with an even number of photons are present, meaning that no event with odd photon number should be measured from an ideal squeezed state.

Two-mode squeezing

In the case where the indices are not equal ($k \neq \ell$) in the term (2) of eq.(2.1.17), we have *two-mode squeezing* (TMS). In this case, considering for simplicity a scenario with only $m = 2$ modes, the operator is of the form

$$\mathcal{S}_{\mathrm{TMS}}(\xi) = \exp\left[\xi \hat{a}_1^{\dagger} \hat{a}_2^{\dagger} - \xi^* \hat{a}_1 \hat{a}_2\right], \tag{2.1.39}$$

corresponding to the mode evolutions

$$\hat{a}_1 \mapsto \hat{a}_1 \cosh r + \hat{a}_2^{\dagger} e^{i\varphi} \sinh r, \tag{2.1.40}$$

$$\hat{a}_2 \mapsto \hat{a}_2 \cosh r - \hat{a}_1^{\dagger} e^{i\varphi} \sinh r, \tag{2.1.41}$$

where again $\xi = r e^{i\varphi}$. The associated 4×4 symplectic matrix is thus given by

$$S_{\mathrm{TMS}} = \begin{bmatrix} \cosh r & 0 & 0 & e^{i\varphi} \sinh r \\ 0 & \cosh r & -e^{i\varphi} \sinh r & 0 \\ 0 & -e^{-i\varphi} \sinh r & \cosh r & 0 \\ e^{-i\varphi} \sinh r & 0 & 0 & \cosh r \end{bmatrix}. \tag{2.1.42}$$

The states generated by the two-mode squeezing symplectic matrix acting on the vacuum are called *vacuum two-mode squeezed states*, and have zero displacement and covariance matrix $\boldsymbol{\sigma}_{\mathrm{TMS}} = \frac{1}{2} S_{\mathrm{TMS}} S_{\mathrm{TMS}}^{\dagger}$. As for the single-mode squeezing, also the two-mode version is a pure state and allows us to write it in terms of two-mode photon states as

$$|\mathrm{TMS}(\xi)\rangle = \frac{1}{\cosh r} \sum_{n=0}^{\infty} (e^{i\varphi} \tanh r)^n |n\rangle_1 |n\rangle_2 . \tag{2.1.43}$$

Once again, also these squeezed states presents only terms with even total number of photons, equally distributed between the two modes. They, therefore, present strong correlations between the photon number between the two modes, i.e. entanglement.

Such entanglement is also manifested in the following important property of TMS states when the two modes are considered singularly. If one of the two modes is traced out, e.g. the second, the state of the other, given by $\rho_1 = \mathrm{Tr}_2(|\mathrm{TMS}(\xi)\rangle\langle\mathrm{TMS}(\xi)|)$, is a thermal state with average photon number $\bar{n} = (\sinh r)^2$.

Some other properties of Gaussian states and transformations

Below we briefly list some additional useful properties of Gaussian states which will be used for the experiments in chapters 4 and 5.

- **Losses.** As already mentioned, passive losses in optical circuits represents the main source of noise in photonics. Their effect on Gaussian states can also be easily described. If the transmittivity of a passive lossy channel is γ (considered uniform over the m modes for simplicity), then the momenta of an input Gaussian state are transformed as [5]

$$\bar{\boldsymbol{x}} \overset{\text{loss}}{\longmapsto} \bar{\boldsymbol{x}}\sqrt{\gamma}, \qquad \boldsymbol{\sigma} \overset{\text{loss}}{\longmapsto} \gamma\boldsymbol{\sigma} + (1-\gamma)\mathbb{1}_{2m}/2. \tag{2.1.44}$$

  The action of loss is therefore to eliminate the off-diagonal terms of the covariance matrix, associated to squeezing, driving the state closer to a diagonal thermal state. As expected, for $\gamma = 0$ the vacuum state is obtained at the output.

- **Tracing out modes.** The partial state of a Gaussian state obtained by tracing out a subset of the modes is still Gaussian, and its momenta can be easily calculated as follows. If in a m mode Gaussian state we want to trace out the k-th mode, the covariance of the reduced $m-1$ mode Gaussian state is obtained by deleting the k-th and $(k+m)$-th rows and columns from the original covariance. Similarly, the displacement of the reduced state is obtained deleting the k-th and $(k+m)$-th elements of the original displacement vector.

2.1.5 *Second-order correlation function*

The different nature of the various quantum states of light discussed in this section is manifested in their different photon statistics. A well-known quantifier to discriminate such statistics is the normalised second-order correlation function:

$$g^{(2)}(\tau) = \frac{\langle \hat{a}_0^\dagger \hat{a}_\tau^\dagger \hat{a}_\tau \hat{a}_0 \rangle}{\langle \hat{a}_0^\dagger \hat{a}_0 \rangle \langle \hat{a}_\tau^\dagger \hat{a}_\tau \rangle}, \qquad g^{(2)}(0) = \frac{\langle (\hat{a}^\dagger)^2 \hat{a}^2 \rangle}{\langle \hat{a}^\dagger \hat{a} \rangle^2} = 1 + \frac{\langle (\Delta \hat{n})^2 \rangle - \bar{n}}{\bar{n}^2}. \tag{2.1.45}$$

This quantity concerns the correlations in the joint detections of two photodetectors, where at the numerator we have coincidences counts, and at the denominator single

counts at each detector. The $g^{(2)}(0)$ thus represents a normalised coincidence count at the two detectors. For an arbitrary single mode input state, such correlation function can be experimentally measured using a beam-splitter and counting at the coincidences of two single-photon detectors at the two outputs; the famous Hanbury-Brown-Twiss interferometer [6]. The values of $g^{(2)}(0)$ for the different classes of quantum states are given by:

- **Coherent states.** Using the properties of coherent states we have
$$g^{(2)}(0) = \frac{\langle(\hat{a}^\dagger)^2\hat{a}^2\rangle}{\langle\hat{a}^\dagger\hat{a}\rangle^2} = \frac{\langle\alpha|\,\alpha^*\hat{n}\alpha\,|\alpha\rangle}{\langle\alpha|\,\alpha^*\alpha\,|\alpha\rangle^2} = \frac{|\alpha|^4}{|\alpha|^4} = 1, \tag{2.1.46}$$
independently from the amplitude α.

- **Single photons.** For a Fock state $|n\rangle$ with a fixed number of photons, there is no uncertainty in the photon number, i.e. $\Delta\hat{n} = 0$, therefore
$$g^{(2)}(0) = 1 - \frac{1}{n}. \tag{2.1.47}$$
In particular, for $n = 1$ the value of $g^{(2)}(0)$ vanishes.

- **Thermal states**. For the Bose-Einstein statistics governing thermal states we have $\langle(\Delta\hat{n})^2\rangle = \bar{n}(\bar{n}+1)$, thus
$$g^{(2)}(0) = 1 + \frac{\bar{n}^2 + \bar{n} - \bar{n}}{\bar{n}^2} = 2 \tag{2.1.48}$$
independently from the average number of photons.

- **Squeezed states**. In the case of a single-mode vacuum squeezed state, with squeezing parameter $\xi = re^{i\varphi}$, we have [7]
$$g^{(2)}(0) = 1 + \frac{\cosh 2r}{\sinh^2 r} > 3, \tag{2.1.49}$$
which is independent from the phase φ, tends to infinity for $r \to 0$ and to 3 for $r \to +\infty$.

2.2 Quantum interference in linear optics

One of the central elements in photonic quantum information processing is quantum interference in linear optical networks. This section is dedicated to give a brief introduction to such effects, with focus on single photon states.

2.2.1 *Hong-Ou-Mandel effect*

We have discussed in the previous section how the behaviour of a single photon injected into a linear optical network is not different from the one of classical coherent light. This is no longer true when more than one photon is involved, due to the possibility of non-classical interference. A simple case of quantum interference is the Hong-Ou-Mandel effect [8], which considers two identical photons injected into two different input ports of a beam-splitter. Coincidence detection is performed at the outputs. Such experimental configuration is represented in Fig. 2.2a.

In a classical scenario, where non-interacting particles are used instead of photons, the global behaviour can be simply described by considering the individual evolution of each particle. In particular, each particle has a equal 50% chance of coming out from the top output port or from the bottom one. This leads to four equiprobable output configurations, with only two (configurations ① and ② in Fig. 2.2a) giving rise to a coincidence detection at the output. The probability to observe the two detectors clicking together is thus $p_{\text{coinc}} = 1/2$.

Now, if we instead consider two indistinguishable single photons, arriving at the same at the beam-splitter, we can write the input two-photon state as $|\psi\rangle_{\text{in}} = |1\rangle_1 |1\rangle_2 = \hat{a}^\dagger_{1,\text{in}} \hat{a}^\dagger_{2,\text{in}} |0\rangle$. The state at the output of the beam-splitter can be obtained by considering the transformation on the mode operators $\hat{\boldsymbol{a}}^\dagger_{\text{out}} = U^\top_{\text{BS}} \hat{\boldsymbol{a}}^\dagger_{\text{in}}$ (see eq.2.1.18), where

$$U_{\text{BS}} = U_{\text{BS}}\left(\frac{1}{\sqrt{2}}\right) = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix} \tag{2.2.1}$$

is the unitary matrix of an ideal $50:50$ beam-splitter. Inverting the mode evolution we can explicitly write the output state as

$$\hat{a}^\dagger_{1,\text{in}} \hat{a}^\dagger_{2,\text{in}} |0\rangle \mapsto \frac{1}{2} \left(\hat{a}^\dagger_{1,\text{out}} + i\hat{a}^\dagger_{2,\text{out}}\right) \left(\hat{a}^\dagger_{2,\text{out}} + i\hat{a}^\dagger_{1,\text{out}}\right) |0\rangle \tag{2.2.2}$$

$$= \frac{1}{2} \Big[\underbrace{i \left(\hat{a}^\dagger_{1,\text{out}}\right)^2}_{①} + \underbrace{i \left(\hat{a}^\dagger_{2,\text{out}}\right)^2}_{②} + \underbrace{\hat{a}^\dagger_{1,\text{out}} \hat{a}^\dagger_{2,\text{out}}}_{③} - \underbrace{\hat{a}^\dagger_{1,\text{out}} \hat{a}^\dagger_{2,\text{out}}}_{④} \Big] |0\rangle \tag{2.2.3}$$

$$= \frac{i}{\sqrt{2}} \left(|2\rangle_1 |0\rangle_2 + |0\rangle_1 |2\rangle_2\right), \tag{2.2.4}$$

where to obtain the last line we have used $|n\rangle = \left(\hat{a}^\dagger_{1,\text{out}}\right)^2 / \sqrt{n!} \, |0\rangle$. Note that in eq.(2.2.3) all possible output configurations shown in Fig. 2.2a are present, but the two configurations with photons in different modes have opposite sign and cancel out. The cases that survive are thus only those where the two photons are grouped together on the same output mode, that is $p_{\text{coinc}} = 0$. No coincidences are observed!

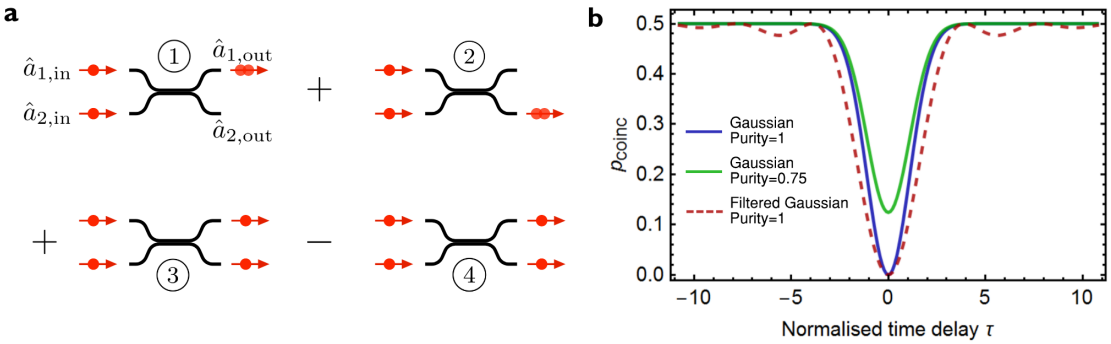


Figure 2.2: Hong-Ou-Mandel experiment. **a,** Experimental configuration of the experiment and possible output configurations. If the experiment is performed with indistinguishable photons, configurations 3 and 4 are identical and with different signs. Therefore, they cancel out, and no coincidence detection is observed at the output. **b,** Examples of Hong-Ou-Mandel dips for photons with different purities and spectral shapes. The reported plots are simulations using eq.(2.2.7).

This is the famous Hong-Ou-Mandel (HOM) effect. The fact that the two photons, which may originate from very distant and uncorrelated sources, must come out together after the beam-splitter makes it "look like" the two photons somehow interact at the beam-splitter. However, what is really happening is not an interaction, but is rather a consequence of the bosonic statistical properties of photons (the HOM effect is in fact a form of Bose-Einstein condensation) arising from the interference of the complex amplitudes associated to multi-photon states, with no classical analog. Nonetheless, the fact that it "looks like" the two photons interact in the HOM effect makes it a very important tool in photonic quantum information processing, as we will see in section 2.2.5.

Hong-Ou-Mandel effect in the general case

In the previous description of the HOM effect we considered for simplicity indistinguishable photons in a pure state arriving simultaneously at the beam-splitter. However, this model typically provides only a very approximate description of real experimental situations. For example, different spectral properties of the two photons and temporal mismatches can add some degree of distinguishability. Additionally, the photon sources may generate them in non pure states.

To include these defects in the model we can first define the *broadband mode operator* to describe the spectral degree of freedom of photons:

$$\hat{A}_k = \int d\omega \psi_k^*(\omega)\hat{a}_k(\omega), \qquad \hat{A}_k^\dagger = \int d\omega \psi_k(\omega)\hat{a}_k^\dagger(\omega), \tag{2.2.5}$$

where $\psi_k(\omega)$ represents the spectral amplitude of a photon, and $\hat{a}_k(\omega)^\dagger$ and $\hat{a}_k(\omega)$ are the creation and annihilation operators for a photon with frequency ω in the k-th spatial mode. A time delay τ on the mode k can be added by replacing $\psi_k(\omega) \mapsto \psi_k(\omega)e^{i\tau\omega}$. Additionally, to include mixedness in the spectral state of the photon, present in realistic photon sources, we can describe the state in each mode of the HOM experiment using density operators

$$\rho_1 = \sum_{i_1} \lambda_{i_1}^2 \hat{A}_{1,i_1}^\dagger |0\rangle \langle 0| \hat{A}_{1,i_1}, \qquad \rho_2 = \sum_{i_2} \mu_{i_2}^2 \hat{A}_{1,j_2}^\dagger |0\rangle \langle 0| \hat{A}_{1,i_2}, \tag{2.2.6}$$

where the indices i_1 and i_2 label the different spectral modes that incoherently contribute in the spectra for the spatial mode 1 and 2, respectively. In this framework, the probability to detect a coincidence as a function of the temporal delay τ between the arriving photons is given by [9]:

$$p_{\text{coinc}}(\tau) = \frac{1}{2} - \frac{1}{2} \sum_{i_1,i_2} \lambda_{i_1}^2 \mu_{i_2}^2 \left| \int d\omega_1 \psi_{1,i_1}^*(\omega) \psi_{2,i_2}(\omega) e^{i\tau\omega} \right|^2 . \tag{2.2.7}$$

The coincidence probabilities for different input states are plotted in Fig. 2.2b as a function of the temporal delay. For large delays, larger than the coherence times of the photons, the two photons become temporally distinguishable and the second integral in eq.(2.2.7) becomes zero. In this case the classical situation is in fact recovered $p_{\text{coinc}}(\tau \to \infty) = 1/2$. When τ is sufficiently small, an HOM dip is observed, with visibility given by

$$\mathcal{V}_{\text{HOM}} = \frac{p_{\text{coinc}}(\tau \to \infty) - p_{\text{coinc}}(0)}{p_{\text{coinc}}(\tau \to \infty)} \tag{2.2.8}$$

$$= \sum_{i_1,i_2} \lambda_{i_1}^2 \mu_{i_2}^2 \left| \int d\omega \ \psi_{1,i_1}^*(\omega) \psi_{2,i_2}(\omega) \right|^2 , \tag{2.2.9}$$

which depends on the ovelap of the spectra of the photons. Note also that, in the case $\rho_1 = \rho_2$, the visibility reduces to

$$\mathcal{V}_{\text{HOM}} = \sum_{i_1} \lambda_{i_1}^4 = \text{Tr}\left(\rho_1^2\right) = \mathcal{P}, \tag{2.2.10}$$

where $\mathcal{P}$ is the purity of the single photon states. The HOM experiment thus also represents a very important diagnostic tool for, e.g., characterising the spectral purity of single photon sources.

2.2.2 *Reversed Hong-Ou-Mandel effect*

In the standard HOM effect the photons start in separate modes, with state $|11\rangle$, and end up in a uniform superposition $(|20\rangle + |02\rangle)/2$ of being both in the top mode and both in the bottom one. Another interesting type of quantum interference to study, which will be extensively used in the experiments reported in this thesis, is the time reversal of such process, known as the *reversed Hong-Ou-Mandel* (rev-HOM) effect. The configuration for rev-HOM experiments is shown in Fig. 2.3a, were we start with the superposition $(|20\rangle + |02\rangle)/2$, apply a phase shift to the bottom mode, combine the two modes on a beam-splitter, and detect coincidences at the two output modes. The action of a phase shifter on a mode operator is simply $\hat{a}_k^\dagger \mapsto e^{i\phi}\hat{a}_k^\dagger$, so that the state injected into the beam-splitter can be written as:

$$|\psi\rangle_{\text{in}} = \frac{1}{\sqrt{2}}\left[(\hat{a}_1^\dagger)^2 + e^{i2\phi}(\hat{a}_2^\dagger)^2\right]|0\rangle\,. \tag{2.2.11}$$

Note that the phase in the superposition state acquires a factor 2, due to the fact that two photons pass through the phase shifter. Applying the BS transformation as before, we arrive, up to a global phase, to the output state

$$|\psi\rangle_{\text{out}} = \cos(\phi)\,|11\rangle - \sin(\phi)\,(|20\rangle - |02\rangle)\,/\sqrt{2}. \tag{2.2.12}$$

The probability to detect a coincidence thus depends on the applied phase as $p_{\text{coinc}} = \cos^2(\phi)$. In the rev-HOM effect quantum interference fringes are thus obtained when varying the phase ϕ, as reported in Fig. 2.3b. Note that the frequency of these fringes is doubled respect to the interference observed from classical light, for which the power in one of the output ports varies as $\propto \cos^2(\phi/2)$.

In the description above we have considered the photons to be identical for simplicity. However, a very important property of the rev-HOM effect is that it remains valid even when the two photons are not indistinguishable. In fact, one can consider the general case where the two photons have different frequencies, say ω_i for one of the two photons, namely the *idler*, and the ω_s for the other one, i.e. the *signal*. We can also include entanglement in the joint spectrum and differences in the spectra associated to the two modes. In this general case, the input state is given by $|\psi\rangle_{\text{in}} = \frac{1}{\sqrt{2}}(|\psi\rangle_1 + |\psi\rangle_2)$, with

$$|\psi\rangle_1 = \int \psi_1(\omega_i,\omega_s)\hat{a}_1^\dagger(\omega_i)\hat{a}_1^\dagger(\omega_s)\,|0\rangle\,, \qquad |\psi\rangle_2 = \int \psi_2(\omega_i,\omega_s)\hat{a}_2^\dagger(\omega_i)\hat{a}_2^\dagger(\omega_s)\,|0\rangle\,. \tag{2.2.13}$$

Here $\psi_1(\omega_i,\omega_s)$ and $\psi_2(\omega_i,\omega_s)$ are the *joint spectral amplitudes* (JSA) associated to each of the two modes (which possess the symmetry $\psi(\omega_1,\omega_2) = \psi(\omega_2,\omega_1)$). In this framework,

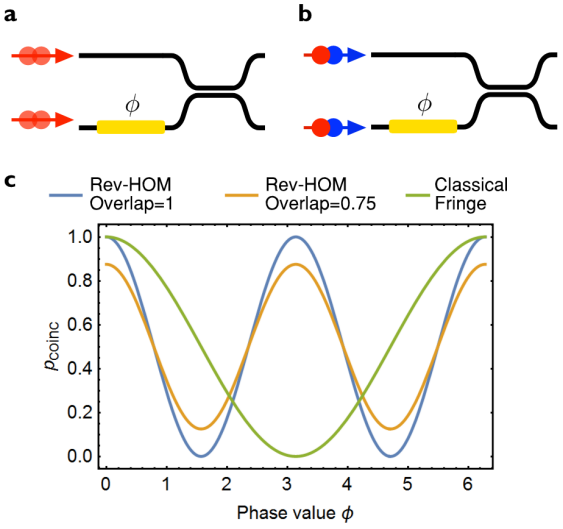


Figure 2.3: Reverse Hong-Ou-Mandel experiments. **a,** Experimental configuration, where a pair of two indistinguishable photons is injected in a superposition of being in two modes. Before the two photon arrive at the beam-splitter, a phase ϕ is inserted on one of the two mode, producing the input state $(|20\rangle + e^{2i\phi}|02\rangle)/\sqrt{2}$. **b,** The quantum rev-HOM interference is observed also if the photons are distinguishble in frequency. **c,** Quantum rev-HOM fringes for different overlaps of the joint spectra, simulated from eq.(2.2.15). The green line represents the classical fringe that would be obtained from the same set-up, whose periodicity is doubled compared to the quantum fringe.

the probability of receiving a coincidence event as a function of the applied phase ϕ is given by (see appendix B for a detailed derivation)

$$p_{\text{coinc}}(\phi) = \frac{1}{2} + \frac{1}{2}\,\text{Re}\left[e^{i2\phi}\int d\omega_1 d\omega_2 \psi_2^*(\omega_1,\omega_2)\psi_1(\omega_1,\omega_2)\right] \tag{2.2.14}$$

$$= \frac{1}{2} + \frac{1}{2}\cos(2\phi+\theta_0)\left|\int d\omega_1 d\omega_2 \psi_2^*(\omega_1,\omega_2)\psi_1(\omega_1,\omega_2)\right|, \tag{2.2.15}$$

where $\theta_0 = \frac{1}{2}\arg\int d\omega_1 d\omega_2 \psi_2^*(\omega_1,\omega_2)\psi_1(\omega_1,\omega_2)$. The visibility of the fringes is exactly given by the overlap between the two joint spectral amplitudes

$$\mathcal{V}_{\text{rev-HOM}} = \left|\int d\omega_1 d\omega_2 \psi_2^*(\omega_1,\omega_2)\psi_1(\omega_1,\omega_2)\right|. \tag{2.2.16}$$

Therefore, the only requirement to have perfect rev-HOM interference is that the joint spectra overlap. However, this can happen even when the photons have different frequencies, and are thus distinguishable! For example, for $\psi_1(\omega_i,\omega_s) = \psi_2(\omega_i,\omega_s) = \delta(\omega_i-\omega_1)\delta(\omega_s-\omega_2)$, with $\omega_1 \neq \omega_2$, the two photons are perfectly distinguishable, but would still provide perfect visibility rev-HOM quantum interference. This highlights that photon indistinguishability is not required to observe interesting quantum effects (the trick here was to start already in a superposition state, even if the objects were distinguishable).

Applications of the rev-HOM effect are, for example, deterministic splitting of degenerate photons [10, 11], quantum metrology [12], and characterisation of photon sources

indistinguishability (i.e. the overlap of their JSAs). The latter application will be extensively used in the experiments reported in this thesis.

2.2.3 *Multi-photon evolution in linear optics*

We have so far described evolutions of few photons in a linear interferometer composed of only one beam-splitter and one phase-shifter. We now describe how we can generalise this to an arbitrary number of input photons in a general linear optical network. As discussed in the previous section, photonic Fock states represent non-Gaussian states of light. As a consequence, the general evolution of Fock states in linear optical circuits cannot be straightforwardly described as in the Gaussian formalism (eq.2.1.34). However, the amplitudes of the output states can be calculated in terms of the *permanent* transfer matrix, which is a matrix function defined as

$$\mathrm{Perm}(A) = \sum_{\sigma \in S_n} \prod_{i=1}^{n} A_{i,\sigma(i)}, \tag{2.2.17}$$

with A being an $n \times n$ matrix and S_n the set of permutations of integers 1 to n. In particular, considering a scenario where an n-photon Fock state $|\boldsymbol{n}\rangle = |n_1 n_2 \dots n_m\rangle$ (with $\sum_{i=1}^{m} n_i = n$) is injected into an m-modes linear optical interferometer with associated $m \times m$ unitary transformation U, the output state can be written in terms of the output Fock states $|\boldsymbol{k}\rangle = |k_1 k_2 \dots k_m\rangle$ (again with $\sum_{j=1}^{m} k_j = n$) as [13]

$$\mathcal{U}(U)\,|\boldsymbol{n}\rangle = \sum_{\boldsymbol{k}:\sum_{j=1}^{m} k_j = n} c_U(\boldsymbol{k}|\boldsymbol{n})\,|\boldsymbol{k}\rangle\,, \tag{2.2.18}$$

where the amplitudes are given by

$$c_U(\boldsymbol{k}|\boldsymbol{n}) = \langle \boldsymbol{k}|\,\mathcal{U}(U)\,|\boldsymbol{n}\rangle = \frac{\mathrm{Perm}\left(U_{\boldsymbol{k},\boldsymbol{n}}\right)}{\sqrt{\left(\prod_{i=1}^{m} n_i!\right)\left(\prod_{j=1}^{m} k_j!\right)}}. \tag{2.2.19}$$

Here, $U_{\boldsymbol{k},\boldsymbol{n}}$ represents the $n \times n$ submatrix of U formed by taking its j-th row k_j times ($i, j \in \{1, \dots, m\}$) and it's i-th column n_i times. These equations allow us to describe the output state of an arbitrary linear optical evolution of photon Fock states. The probability to detect the configuration $\boldsymbol{k}$ at the output is given by

$$p_U(\boldsymbol{k}|\boldsymbol{n}) = |\,\langle \boldsymbol{k}|\,\mathcal{U}(U)\,|\boldsymbol{n}\rangle\,|^2 = \frac{|\mathrm{Perm}\left(U_{\boldsymbol{k},\boldsymbol{n}}\right)|^2}{\left(\prod_{i=1}^{m} n_i!\right)\left(\prod_{j=1}^{m} k_j!\right)}. \tag{2.2.20}$$

In particular, for $n = 1$ we recover the probability that a photon injected in the i-th mode emerges on the j-th, given by $p(j|i) = |U_{ji}|^2$, as we already saw in section 2.1.3.

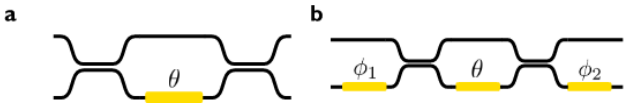


Figure 2.4: Linear optical transformations on two modes. **a,** The Mach-Zehnder interferometer. **b,** Circuit to implement arbitrary unitary operations on two modes.

The association between the permanent function and the linear evolution of multiple bosons was first observed by Caianiello [14]. An important result about permanents, due to Valiant [15], is that the classical complexity for their calculation is in the $\#P$-hard class, which makes the evaluation intractable on classical machines for sufficiently large matrices. The connection between the permanents and amplitudes in linear optical circuits was given by Scheel [13], who also suggested that it would imply that simulating linear optical experiment is $\#P$-hard. This suggestion was made formal by Aaronson and Arkhipov, who showed that even the approximate simulation is $\#P$-hard [16]. This result propelled linear optical protocols that aim at showing near-term quantum advantages, such as the *boson sampling* protocols that will be discussed in chapter 4.

2.2.4 *Linear optical circuits*

Photonic quantum information processing consists largely of designing interferometers to perform a desired evolution. Fortunately, there exist simple schemes that allow us to decompose any unitary evolution on an arbitrary number of modes using only cascaded beam-splitters and phase-shifters. In this way, in photonics, once we find the unitary U required for a particular protocol, we can immediately derive a physical circuit that allows us to implement it.

Mach-Zehnder interferometer

We start considering the simplest case, i.e. arbitrary unitary operations on two modes. This allows us to introduce a key element in linear optics: the *Mach-Zehnder interferometer* (MZI). The circuit for a MZI, shown in Fig. 2.4a, comprises a phase-shifter embedded between two beam-splitters. The resulting unitary operation is given by

$$U_{\mathrm{MZI}}(\theta) = U_{\mathrm{BS}} U_{\mathrm{PS}}(\theta) U_{\mathrm{BS}} = i e^{-\frac{i\theta}{2}} \begin{bmatrix} \sin(\theta/2) & \cos(\theta/2) \\ \cos(\theta/2) & -\sin(\theta/2) \end{bmatrix}. \tag{2.2.21}$$

Note that, up to phases, the MZI unitary represents the matrix of a beam-splitter with an arbitrary reflectivity $\eta = \sin^2(\theta/2)$. An MZI with two phase-shifters at each side (see Fig. 2.4b) allows us to implement arbitrary operations on the SU(2) group. To see this, we can recall that any SU(2) matrix can be written in terms of Pauli rotations as [17]

$$U = R_Z(\gamma) R_X(\beta) R_Z(\alpha), \tag{2.2.22}$$

where $R_X(\phi) = e^{i\phi\hat{\sigma}_z/2} = U_{\rm PS}(\phi)$ and

$$R_X(\theta) = e^{i\theta\hat{\sigma}_x/2} = -ie^{\frac{i\theta}{2}} U_{\rm PS}(\pi/2) U_{\rm MZI}(\theta) U_{\rm PS}(\pi/2), \tag{2.2.23}$$

which gives, up to global phases, $U = U_{\rm PS}(\phi_2) U_{\rm MZI}(\theta) U_{\rm PS}(\phi_1)$ for $\phi_1 = \alpha + \pi/2$ and $\phi_2 = \gamma + \pi/2$.

Universal linear optics decompositions

We can now proceed to describe schemes generalising the MZI to universal circuits on an arbitrarily large number of modes. The building block for such schemes are 2×2 operations composed of a MZI and a phase-shifter (see inset of Fig. 2.5), which corresponds to the unitary matrix

$$T_k(\theta, \phi) = U_{\rm PS}(\phi) U_{\rm MZI}(\theta) = ie^{-\frac{i\theta}{2}} \begin{bmatrix} e^{i\phi}\sin(\theta/2) & e^{i\phi}\cos(\theta/2) \\ \cos(\theta/2) & -\sin(\theta/2) \end{bmatrix}, \tag{2.2.24}$$

called *Givens* rotations. Here the index k indicates that the MZI is mixing the k-th and $k+1$-th adjacent modes, and is acting as identity on all the others. The goal is to decompose an arbitrary $m \times m$ unitary matrix U into a series of 2×2 matrices $T_k(\theta, \phi)$. To achieved this, we can use schemes based on the fact that applying the matrix $T_k(\theta, \phi)$ to the right of U, i.e. taking $UT_k(\theta, \phi)$, corresponds to mixing the k-th and $k+1$-th columns of U. Similarly, applying $T_k(\theta, \phi)$ to the left, i.e. $T_k(\theta, \phi)U$ corresponds to mixing the k-th and $k+1$-th rows of U. The general idea is then to proceed iteratively, where at each step two consecutive modes are mixed applying a Givens rotation choosing the values (θ, ϕ) so that zeros are generated in the resulting matrix. At the end of such procedure, i.e. when all off-diagonal elements are zero, we are left with a diagonal unitary matrix. Inverting the resulting matrix equation we can finally obtain a decomposition of the form

$$U = D \prod_{i \in \mathcal{P}} T_{k_i}(\theta_i, \phi_i), \tag{2.2.25}$$

for some pattern $\mathcal{P}$ which depends on the scheme adopted. This equation represents a network of Givens rotations and final phase-shifters to implement the last diagonal matrix. The number of Givens rotations, i.e. of MZIs, required depends on the number of iterations needed to delete all off-diagonal elements. As in general the number of independent off-diagonal elements in a $m \times m$ matrix is $m(m-1)/2$, the number of MZIs in the decomposition will be

$$\#_{\rm MZI} = \frac{m(m-1)}{2} \tag{2.2.26}$$

independently from the scheme adopted.

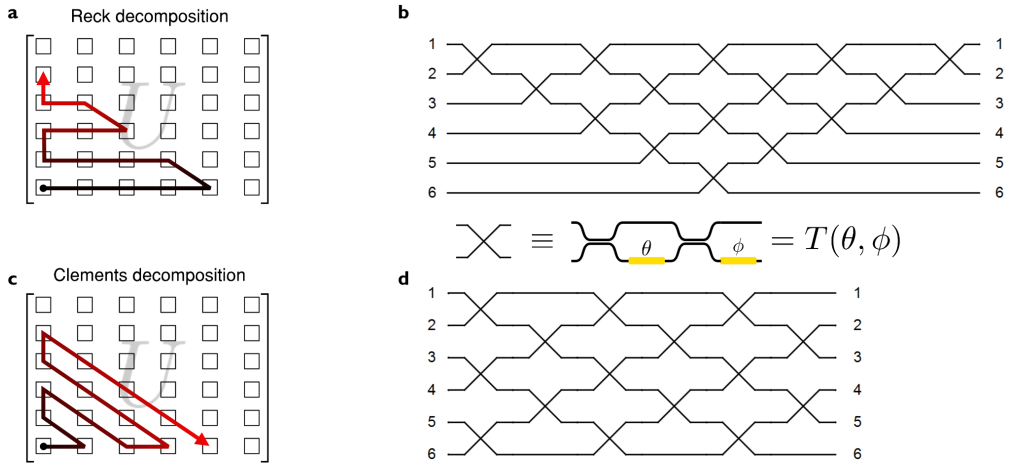


Figure 2.5: Universal linear optics schemes. **a,** Elimination pattern for the Reck scheme. The resulting structure of this decomposition is a triangular network of MZIs, shown in **b** for an exemplary 6-mode case. **c,** Elimination pattern for the Clements decomposition, which results in the square network pictured in **d**, optimal in terms of circuit compactness and symmetry. Inset: each of the crossed lines in the scheme representations is a MZI with an external phase-shifter, implementing a Givens matrix.

The chronologically first decomposition developed for universal linear optics is the *Reck scheme*. Although exactly the same decomposition was known and adopted in the microwave beam forming community since 1965, known as *Nolan synthesis* [18, 19], it was only in 1994 that Reck et al. introduced it in the context of quantum photonics [20]. In this scheme, the procedure for cancelling the matrix elements proceeds row-by-row, as shown in Fig. 2.5a, where at each step the values of the angles in the Givens rotations are iteratively calculated. The decomposition results in a triangular network of $m(m-1)/2$ MZIs (see in Fig. 2.5b).

A second scheme for universal linear optics, recently proposed by Clements et al. [21], results instead in a square network of MZIs, as shown in Fig. 2.5c. Instead of following the rows, the matrix elements elimination pattern in this scheme proceeds by diagonals (see in Fig. 2.5d). Again, all phases in the Givens rotations can be algorithmically inferred for an arbitrary unitary in an iterative fashion. Apart from its increase compactness, the main advantage of the Clements scheme respect to the Reck decomposition is that the optical depth, i.e. the number of elements a photon passes through, is independent on the input and output mode of the photons. In particular, photons always pass through m MZIs in a m-mode Clements scheme. In contrast, if a photon is injected and collected on the top mode of a Reck scheme, it would only see a single MZI, while it will have to pass through m MZIs to come out from the bottom port. This property makes the Clements

scheme more resilient to errors due to losses. On the other hand, both Reck and Clements schemes require the same number of total optical components, and have been shown to be equivalent in terms of robustness to noises in setting the phase shifters [22, 23].

As the number of modes starts to increase, the number of optical components that need to be interconnected in practical implementations of universal linear optical schemes start to be very daunting. A particular challenging issue is given by the requirement of extremely high phase stability inside such large interferometers to achieve sufficiently high fidelities. For these reasons, multi-mode universal linear optical circuits have not been experimentally reported in bulk or fibre optics. Only with the advent of integrated photonics, where intrinsically stable interferometers can be build on photonic chips, universal linear optical circuits have become practically relevant, rather than a theoretical curiosity. The first demonstration of a reprogrammable universal multi-mode interferometer has been achieved in 2015 on a 6-modes silica integrated device [24] (see also Fig. 5.3b).

2.2.5 *Quantum computing with linear optics*

Virtually all quantum states of light we described in section 2.1 can be used to encode quantum information and process it in a digital manner (i.e. with well-specified gates on qubits). Protocols based on Gaussian states, such as coherent and squeezed states, are usually referred to as *continuous-variables quantum computing*, while those based on single photons are called *discrete-variables quantum computing*. In this thesis we will be mostly concerned with the latter, although we will also make use of Gaussian states in the experiments reported in chapters 4 and 5.

Encoding qubits and qudits with photons

In photonics a qubit is typically encoded in the Fock state of a single photon in two optical modes. Although these modes can represent many different photonics degrees of freedom (e.g. polarisation, spectrum, time, orbital angular momentum, etc.), here we will consider the path of the photon. The qubit is then encoded in a single photon as pictured in Fig. 2.6, where the mapping between the logical state of the qubit and the Fock state of the photon is

Logical state		**Fock state**	
$\|0\rangle$	$\longleftrightarrow$	$\|1\rangle_0 \|0\rangle_1$	(2.2.27)
$\|1\rangle$	$\longleftrightarrow$	$\|0\rangle_0 \|1\rangle_1$	(2.2.28)

We therefore say that the qubit is in the logical state $|i\rangle$ if the single photon occupies its i-th mode. This definition can be straightforwardly generalised to encode d-dimensional

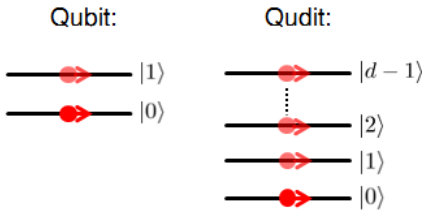


Figure 2.6: Encoding of a qubit (left) and a qudit (right) on the spatial modes of a single photon.

qudits in case the photon can occupy $m = d$ different optical modes. An arbitrary single-qubit operation on a photon then correspond to an arbitrary unitary U on two modes, which, as we have seen, can be simply implemented via a MZI. The analog for qudits are arbitrary unitaries on d-modes, which can be implemented via universal linear optical schemes.

Probabilistic two-photon gates in linear optics

The problems for photonic quantum computing start when it comes to performing entangling two-qubit operations, made difficult by the absence of photon-photon interaction. This lack of interaction is due to the linearity of the transformation eq.(2.1.16), which governs all the optical processes described so far. Without non-linearities, the task of generating entanglement in photonics is quite daunting. Nonetheless, there are ways to avoid this issue. We will now see that HOM interference combined with projective measurements can still make the photons act "as if" they had interacted, even if they don't. The drawback of using this type of disguised interaction, rather than a real one, is that the implementation of two-qubit gates will be intrinsically probabilistic.

To see this, we describe the photonic implementation of a `C-Phase` $= \mathrm{diag}(1, 1, 1, -1)$ two-qubit gate via the interferometer shown in Fig. 2.7, developed by Ralph et al. and, independently, by Hofmann and Takeuchi [25, 26]. In this scheme the interferometer has 6 modes, and consists of three beam-splitters with transmittivity 1/3. The two input qubits are encoded in two photons entering the four central modes of the interferometer; the top two central modes associated to the first qubit, and the second two associated to the second qubit. Crucially, we keep only the output cases with one photon in the top two central modes, and the other one in the bottom two central modes, i.e. when the qubits are still well-defined also in the output. All other possible output configurations are discarded. This selective process, referred to as *post-selection*, is frequently used in quantum photonics. It is easy to calculate that, independently from the input states of the two qubits, the probability of ending up in a permitted output configuration is 1/9,

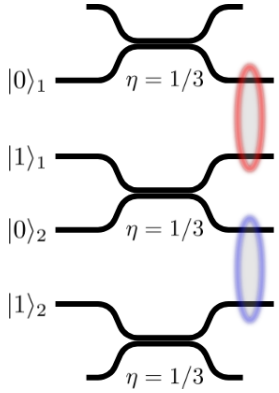


Figure 2.7: Schematic of the 6-mode interferometer performing a photonic probabilistic `C-Phase` entangling gate with linear optics using post-selection. The photons encoding the two qubits and injected in the four central modes. The successful action of the gate is conditional on detecting one photons on the top two central output modes (red), and one on the bottom two central modes (blue). All other outputs configurations are discarded. The probability of success for this gate is 1/9.

and, whenever that happens, the computational basis states are coherently mapped into

$$|00\rangle \mapsto \frac{1}{3}|00\rangle \qquad |01\rangle \mapsto \frac{1}{3}|01\rangle \tag{2.2.29}$$

$$|10\rangle \mapsto -\frac{1}{3}|10\rangle \qquad |11\rangle \mapsto \frac{1}{3}|11\rangle \tag{2.2.30}$$

This transformations correspond to the gate matrix $U = \frac{1}{3}\mathrm{diag}(1, 1, -1, 1)$, which is the desired `C-Phase` gate up to a relabel of the output modes and the prefactor $\frac{1}{3}$ (due to having the operation succeeding with a probability 1/9). More efficient ways to implement arbitrary controlled operations in photonics will be discussed in chapter 6.

Universal linear optical quantum computing

In the linear optical `C-Phase` scheme described above we can identify two key elements: quantum interference in the linear optical circuits, and post-selection. Note that the post-selection is effectively a projective measurement, which projects the state in the subspace of the Hilbert space spanned by the accepted configurations. A crucial point here is that measurements represent non-unitary operations, and their action can thus be tailored to be equivalent to a non-linearity between the two photons, which mediates an interaction. Such effective interaction is often called *measurement-induced non-linearity*. Post-selection is therefore a fundamental ingredient in linear optical quantum computing, as it allows us, quite surprisingly, to introduce non-linearities between non-interacting particles.

A drawback is that post-selection is intrinsically probabilistic: it works only when the measurement projection (governed by Born's rule) is successful. Moreover, post-selected gates cannot be cascaded, as that could bring us to accept final configurations arising from forbidden states after the first gate. A possible idea to avoid this type of situation

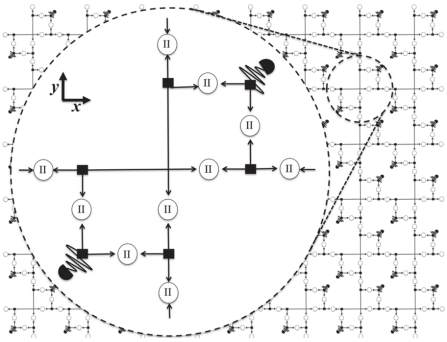


Figure 2.8: Schematic for the modular photonic MBQC architecture reported in Ref. [27]. The unit cell of the architecture is zoomed in, where the black squares represents sources of three photon GHZ states, and circles represent heralded probabilistic two-photon gates (type-II fusion gates). Image reproduced from Ref. [27].

in linear optical quantum computing is to add additional ancillary photons which herald the success of a gate. The photons are then stored and passed to the following stage of the computation only when all previous operations on all photons of the computer have succeeded. This type of protocol for the photonic implementation of two-qubit gates based on feed-forward techniques, combined with single photon operations, can be in principle used to perform universal quantum computation.

The first to realise that, although lacking photon-photon interaction, linear optics is compatible with scalable universal quantum computation were Knill, Laflamme and Milburn (KLM) in 2001 [28]. In their seminal paper they showed that it is, in fact, possible to induce optical non-linearities by using measurement and feed-forward, thus proving that linear optical quantum computing is, in principle, achievable. However, the resource requirement for a practical realisation of a KLM optical quantum computer, although polynomial, are still very prohibitive. Fortunately, remarkable advances in theoretical architectures for linear optical fault-tolerant universal quantum computation have been developed in the last few years. State-of-the-art approaches are now based on the on the *measurement-based quantum computing* (MBQC) paradigm put forward by Raussendorf and Briegel [29, 30]. In contrast to the standard gate-based model of quantum computing, MBQC is based on performing single qubit measurements on a initial highly entangled resource state, known as a cluster state [31], in order to process quantum information. Recent proposals based on linear optical MBQC, mainly developed by Browne and Rudolph, made a number of significant improvements in decreasing the required linear optical resources needed to generate large photonic cluster states, removing the need of quantum memories (only passive delay lines are needed) and decreasing the stability requirements [32]. An exemplary schematic of modern architectures for photonic MBQC is pictured in Fig 2.8. It can be observed that this type of linear optical quantum computation is very modular, and each photon only has to pass through a limited number of optical components. Subsequent theoretical improvements to the

scheme have also enabled fault-tolerant computation in the presence of up to 50% loss per photon using ideas from percolation theory [27, 33, 34]. These improvements, together with the technological ones, are making linear optical quantum computing a realistic prospect [27].

The rest of this thesis is in large part dedicated to progress towards the development of a technology able to match the requirements for these linear optical architectures.

2.3 Integrated quantum optics

One of the features that stands out when observing the circuits previously described in this chapter is the remarkable level of complexity required for linear optical quantum applications compared to standard optical interferometers, such as those employed in, Young's, Michelson's, or Mach-Zehnder experiments. For instance, the number of beam-splitters and phase shifters required to implement the circuits in Fig. 2.5 and Fig. 2.8 are orders of magnitude higher. At the same time, in such interferometers, phase stability needs to be maintained between the thousands of possible optical paths that photons can undertake. It is evident that standard approaches developed in the last centuries, such as bulk optics and fibre optics, where photons propagate in free space between forests of mirrors or in long spools of single-mode fibres, cannot represent viable solutions to build stable and controllable interferometers at this scale.

Fortunately, as we will see in the remaining of the thesis, these issues can be addressed via the integrated quantum photonics approach. Integrated photonics represents a technology where the essential components required to generate, manipulate, and measure light are put together on a common platform. The field started in the late 1960s, when the first demonstrations of light propagating in waveguides embedded in standard optical materials, such as dielectric organic films and glasses, were reported [35, 36]. Thanks to the development of more refined fabrication techniques, integrated photonic circuits can now be fabricated in a wide variety of materials, each providing different properties. These include lithium-niobate, where fast modulators can be embedded, III-V semiconductors (e.g. indium-phospide) allowing laser sources and electro-optics, and low loss linear circuits in silica. Since the mid-90s, Silicon has emerged as one of the prominent materials for integrated optics, in particular the silicon-on-insulator (SOI) technology. In this platform, due to the strong refractive indices difference with the substrate, high-confinement silicon waveguides and high density circuits with intrinsic phase stability are achievable. Moreover, SOI fabrication is compatible with the CMOS technology used for silicon electronics, meaning that scalable fabrication of the photonic circuits is possible. Silicon photonic devices are now readily available from commercial foundries.

Starting in 2008, also the quantum optics community has started moving toward inte-

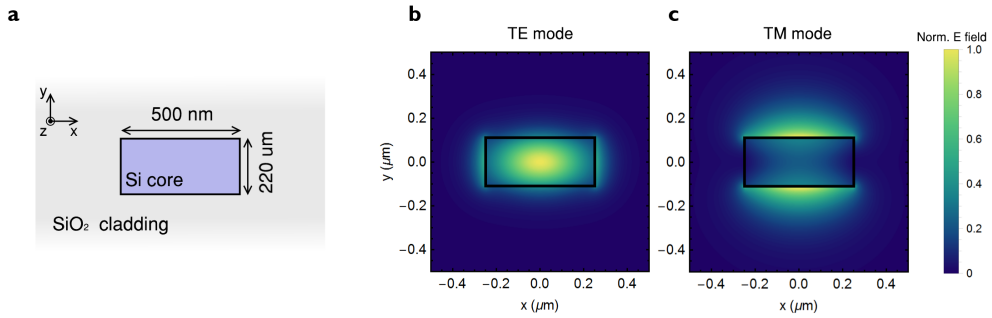


Figure 2.9: Integrated waveguides. **a,** Cross section of a typical strip waveguide in silicon. The associated distribution of the transverse field is shown in **b** and **c** for the lowest order TE_0 and TM_0 modes of the waveguide, respectively.

grated approaches [37]. Currently, the most developed platforms for quantum applications include directly written glass waveguides, silica, silicon-nitride, lithium-niobate, etc. In this thesis, we will focus on developing large-scale quantum systems using semi-conductor photonic devices in silicon.

2.3.1 *Standard components for linear optics in silicon*

We now briefly review the main building blocks for silicon photonics that will be used to implement the large-scale integrated quantum photonics circuits presented in next chapters. We limit the description to the elements used in the experiments reported in this thesis, while we omit components (e.g. micro-resonators and carrier-depletion phase modulators) which, although relevant to integrated quantum optics, have not been used here. The interested reader can refer to [38, 39] for a more exhaustive list.

Waveguides

Waveguides transport optical signals between different regions of an integrated photonic network, and are thus one of the most important elements in optical circuits. They essentially consist of a core region of a material with refractive index n_{core} embedded in a cladding of smaller refracting index $n_{\text{clad}} < n_{\text{core}}$. Similarly as in optical fibers, light is confined inside the core by total internal reflection. Here we will focus on *strip waveguides*, shown in Fig. 2.9, where the core is fully embedded in the cladding and the confinement is maximised. Light travels along the waveguide (z direction) according to Maxwell's equations. The transverse field can be described in terms of discrete modes of

electric field profiles $A(x, y)$, given by the solutions of the Helmoltz equation [39]:

$$\left[\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} - \frac{\omega^2}{c^2} n^2(x, y)\right] A_m(x, y) = -\frac{\omega^2}{c^2} n^2_{\text{eff},m} A_m(x, y), \tag{2.3.1}$$

where m indexes the different modes (i.e. the eigenfunctions of the equation), $n_{\text{eff},m}$ is the *effective refractive index* for the mode $A_m(x, y)$ (associated to its eigenvalue), and ω is the frequency of the propagating light. For a rectangular waveguide, we can also define quasi-TE and quasi-TM modes as those having the electric field principally polarised along the x and y axis, respectively. The propagation along the waveguide axis z for the m-th mode is described by an evolution $e^{-i\beta_m z}$, where $\beta_m = \frac{\omega}{c} n_{\text{eff},m}$ is the wave-vector. The associated group velocity for the propagating mode m is then given by $v_{g,m} = c(n_{\text{eff},m} - \lambda \frac{d}{d\lambda} n_{\text{eff},m})^{-1}$, where we have included the dependence of the effective refractive index on the wavelength.

In this thesis we will consider only *single-mode* waveguides, i.e. supporting only the zero-th order modes of transmission (TE_0 and TM_0). We show in Fig. 2.9 an example of these modes for a typical SOI single-mode rectangular waveguide, obtained via a numerical eigenmode solver (Lumerical). Losses in the propagation through integrated waveguides are mostly due to roughness at the waveguide facet between the core and the cladding, arising due to imperfect fabrications. This implies that modes with field profiles of higher intensities closer to the waveguide facet experience higher losses. For this reason, we will always avoid using light polarised in the TM_0 mode, which, as can be seen from Fig. 2.9, is highly localised at the borders of the core and is thus much more lossy than the TE_0. Silicon is also not transparent in the visible or near infrared region, but it is transparent to wavelengths longer than a micron. Silicon quantum photonics experiments reported in this thesis are therefore all performed at $\simeq 1550$ nm, which is inside the so-called C telecommunication band. Typical losses experienced in standard SOI waveguides (as the one in Fig. 2.9) at these wavelengths are approximately -3 dB/cm with current fabrication technologies.

Integrated beam-splitters

As discussed in previous sections, one of the two main building blocks in linear optics is the beam-splitter. In integrated optics there are two main approaches to implement it: *directional couplers* and *multi-mode coupler*. The directional-coupler component consists in two waveguides brought together for a length z (see Fig. 2.10), called *interaction region*. If the distance δx between the two waveguides in the interaction region is small enough, the fields associated to the two waveguides become evanescently coupled in the cladding, enabling light to tunnel between the two optical modes. The evolution of

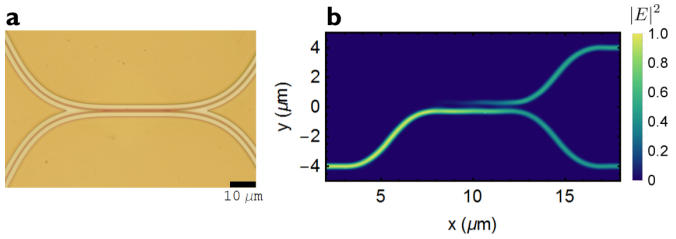


Figure 2.10: Directional coupler. **a,** Optical microscope image of a directional coupler in silicon. **b,** FDTD simulation of a 50:50 directional coupler implemented using coupled waveguides in silicon, showing the evolution of the field in the coupling region.

the field amplitudes in such a structure, simulated via the finite-difference time-domain method (FDTD), is shown in Fig. 2.10b. A formal description of such evolution is given by the coupled mode equation [39, 40]:

$$\frac{\partial A_1(z)}{\partial z} = -i\kappa A_2(z) \tag{2.3.2}$$

$$\frac{\partial A_2(z)}{\partial z} = -i\kappa A_1(z), \tag{2.3.3}$$

where $A_1(z)$ and $A_2(z)$ represent the field amplitude of each of the two modes as a function of the interaction length, and $\kappa \propto e^{-\alpha \cdot \delta x}$ is the coupling constant (α is here a material-dependent constant). The solution to this linear differential equation can be written as

$$\begin{bmatrix} A_1(z) \\ A_2(z) \end{bmatrix} = \begin{bmatrix} \cos(\kappa z) & -i\sin(\kappa z) \\ -i\sin(\kappa z) & \cos(\kappa z) \end{bmatrix} \begin{bmatrix} A_1(0) \\ A_2(0) \end{bmatrix} = U_{\mathrm{DC}}(z) \begin{bmatrix} A_1(0) \\ A_2(0) \end{bmatrix}. \tag{2.3.4}$$

To implement a $50:50$ beam-splitter we have to choose an interaction length $z = \pi/4\kappa$, in which case the unitary transfer matrix of the directional coupler is

$$U_{\mathrm{DC}}(\frac{\pi}{4\kappa}) = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -i \\ -i & 1 \end{bmatrix}, \tag{2.3.5}$$

which corresponds, up to a phase, to the desired 2×2 unitary of a balanced beam-splitter $U_{\mathrm{BS}}(\frac{1}{2})$, see eq.(2.1.20). Directional couplers are very low loss components, which essentially just amounts to the standard transmission loss in the coupling region (typically less than 0.05 dB can be achieved, depending on the quality of the waveguides). On the other hand, they are very sensitive to small errors in the coupling distance. As a consequence, fabricating perfectly balanced beam-splitters is very difficult in the silicon platform. Using the same principle of the directional coupler, one can think of building low-loss interferometers by coupling together multiple waveguides. Such structures form *continuous random walks*, and will be experimentally implemented in chapter 4.

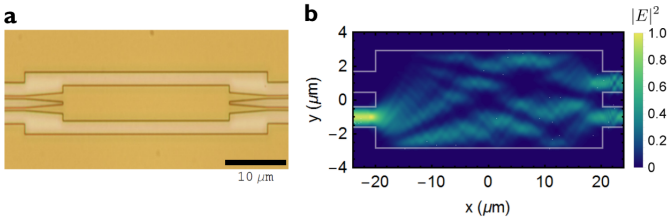


Figure 2.11: Multi-mode interferometer. **a,** Optical microscope image of an MMI in silicon. **b,** FDTD simulation of a 2×2 MMI, showing the evolution of the field inside the multi-mode region, and the collection in the point where the interference forms two specular images.

Compared to directional couplers, designing balanced beam-splitters via a multi-mode interferometer (MMI) has the advantage of being much more robust to fabrication errors [41]. A schematic of the MMI component is shown in Fig. 2.11. The general idea is to inject the two input single-mode waveguides into a larger multi-mode region. In this transition, the input modes are coupled to the multiple higher-order modes supported in the large waveguide. As these modes propagate with different group velocities inside the multi-mode waveguide, they accumulate different phases, giving rise to interference patterns, as shown in the FDTD simulation in Fig. 2.11b. Due to beatings in this interference, repeated images of the input plane are formed at periodic lengths. By an appropriate tailoring of the length of the expanded section, we can collect the output light exactly when the image created by each input is equally split between the two outputs (see Fig. 2.11b). In such case, the transformation operated by the MMI is exactly that of a balanced beam-splitter, given by the transfer matrix

$$U_{\text{MMI}} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix}. \tag{2.3.6}$$

While such transformation is not significantly affected by small errors in the dimension of the MMI, implying fabrication robustness, this component typically presents slightly higher losses than the directional coupler. Most of the integrated beam-splitters used in the experiments presented in this thesis are based on the MMI design, with typical observed losses of $\simeq -0.1$ dB per MMI.

Waveguide crossers

In the fabrication of silicon photonics devices, the photonic components are typically fabricated on a single silicon layer. This means that all circuits are necessarily planar. In such two-dimensional circuits, an important element is the *waveguide* crosser, which allows to swap waveguides. For the experiments reported in this thesis, all the crossers were designs using multi-mode structures similar to the ones used for the MMI [42]. The idea here is to cross the multi-mode waveguides in the exact point where the first image of the input is created. In that point the two modes associated to the two inputs are

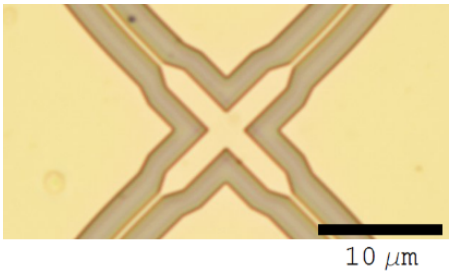


Figure 2.12: Optical microscope image of a waveguide crosser in silicon based on multi-mode interferometers.

orthogonal, thus avoiding cross-talk. The loss in this structures can be kept quite low, typically less than -0.1 dB per crosser.

Integrated phase-shifters

The remaining building block for linear optical interferometers is the integrated phase-shifter. In all experiments reported in this thesis, phase shifters have been implemented exploiting the thermo-optic effect in silicon. In particular, to control the phase in a waveguide, a resistive metal film is designed to be on top of it. Electrical control of the voltage applied to this resistance allows us to tune the amount of power thermally dissipated in it, and hence the temperature of the waveguide underneath the metallic film. Via the thermo-optic effect, a change of temperature induces a change Δn in the refractive index in the waveguide, which in turn induces a phase-shift

$$\Delta\phi = \frac{2\pi L}{\lambda}\Delta n = \frac{2\pi L}{\lambda}\frac{dn_{\text{eff}}}{dT}\Delta T. \tag{2.3.7}$$

The phase can thus be precisely tuned by controlling the voltage applied to the resistence [43]. These components are commonly called *thermal phase-shifters*. In typical silicon waveguides, a change in temperature as high as $\Delta T \simeq 100$K is required to obtain 2π phase shift, corresponding to a dissipated power of approximately 50 mW on the heater. In circuits comprising several tens of phase shifters, as in the experiment reported in chapter 3, dissipating all the heat generated by the heaters in a mm-scale silicon chip becomes a significant issue.

Asymmetric Mach-Zehnder on-chip filters

Another class of elements which will be extensively used in the experiments reported in this thesis are on-chip filters. These will be used to separate photons with different frequencies, or to suppress the bright pump light in the chip after the photon generation. While resonating structures, such as micro-rings, represent viable designs [38], here we will

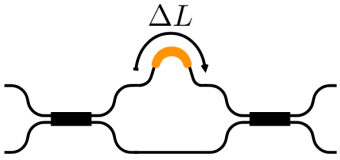


Figure 2.13: Schematic of the on-chip filter structure based on an asymmetric Mach-Zehnder interferometer. Due to a path mismatch ΔL, different relative phases are acquired inside the MZI which depend on the wavelength of the photons. Photons that experience constructive or destructive interference are then separated, obtaining a filter.

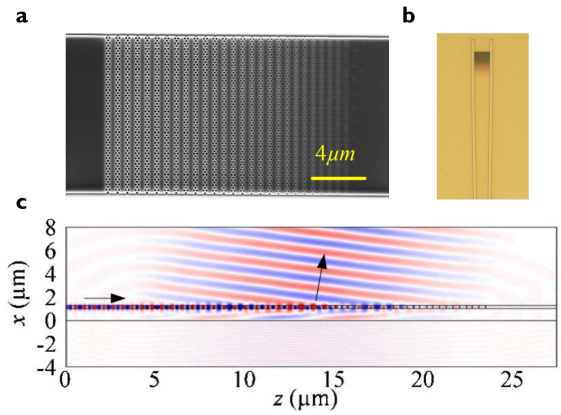


Figure 2.14: Grating couplers for coupling light in and out of silicon chips. **a,** Scanning electron microscope image of the photonic crystal structure in an apodised grating coupler used to direct light vertically out of the waveguide plane and into a single mode fibre. **b,** Optical microscope image of a grating coupler, showing the initial part of a 400μm-long taper used to adiabatically convert the waveguide mode from the size of a single-mode fibre to the one of a single-mode waveguide in silicon. **c,** (Image reported in Ref. [44]) FDTD simulation of the light being vertically scattered out of the waveguide through the grating coupler.

focus on a simpler structure based on *asymmetric Mach-Zehnder* interferometers (AMZIs). This element, pictured in Fig. 2.13, is based on the idea that, if in a Mach-Zehnder one of the two arms is longer than the other by a length ΔL, then the phase acquired in the two arms will be different by a factor

$$\Delta\phi = 2\pi\frac{\omega\Delta L}{v_g}, \tag{2.3.8}$$

where v_g is the group velocity in the single-mode waveguide. The phase, and hence the interference, depends linearly on the frequency. Therefore, to separate photons whose frequencies differ by $\Delta\omega$, we can choose a path difference $\Delta L = v_g/\Delta\omega$ so that constructive interference is obtained for the first photon ($\Delta\phi(\omega_0) = 0$), which will emerge from the top port of the MZI. On the other hand, for the second photon $\Delta\phi(\omega_0+\Delta\omega) = \pi$, and it will thus exit on the bottom output due to destructive interference. We therefore have a filter, with free spectral range FSR $= v_g/\Delta L$. A phase shifter inside the AMZI also enables to rigidly shift the filtered spectrum, allowing us to precisely tune the central frequency ω_0.

Grating couplers to fibre

The final component we discuss is the *grating coupler*, which allows us to couple light in and out of the chip via single-mode optical fibres. We will use this element in all the

experiments reported in thesis to inject pump light into the chip, and collect photons out of the chip for detection. A significant problem one has to face when coupling light between a single-mode fibre and a silicon chip is that the size of the circular core in an optical fibre is 8 μm, while silicon waveguides are less than a μm wide. To achieve an efficient coupling between these two modes with different sizes, the grating coupler structure first uses a taper to adiabatically change the mode size in the waveguide to match it with the one of the fiber, as shown in Fig. 2.14b. Then, it uses constructive interference between distributed scatterers (see Fig. 2.14a) to direct light out of the waveguide plane, and straight into a fibre mode, as can be observed from Fig. 2.14c. This vertical scattering of the light is achieved via periodically placed teeth, whose period determines the direction of the coupling (which differs depending on the used wavelength) [45]. Typical losses for grating couplers are approximately -3 dB. For the experiment reported in chapter 4, however, we will use a special design for ultra-low-loss grating coupling with losses below -1 dB [44, 46].

2.4 Generation of quantum states of light in integrated sources

Having presented the required components to manipulate photons on a chip, we can pass to describe integrated sources of quantum states of light. In particular, we will discuss here sources based on two-mode or single-mode squeezed vacuum. These states can be generated in stimulated non-linear processes arising from interaction of a pump light with a medium [2].

2.4.1 *Spontaneous four-wave mixing*

Interactions between light and a medium can be described as a non-linear dependence of the polarisation field in the medium $\boldsymbol{P}$ from the electric field $\boldsymbol{E}$:

$$P_i(\boldsymbol{r},t) = \epsilon_0 \left[\chi^{(1)}_{ij} E_j(\boldsymbol{r},t) + \chi^{(2)}_{ijk} E_j(\boldsymbol{r},t)E_k(\boldsymbol{r},t) + \chi^{(3)}_{ijk\ell} E_j(\boldsymbol{r},t)E_k(\boldsymbol{r},t)E_\ell(\boldsymbol{r},t) + \dots\right], \tag{2.4.1}$$

where ϵ_0 is the vacuum permittivity and $i,j,k,\ell \in \{x,y,z\}$. The tensor $\chi^{(n)}$ represent the susceptibility of order n (sums are implied over the indices of the electric fields). In quantum optics, the most used types of non-linear effects are probably second-order processes dictated by the $\chi^{(2)}$ of the material. These include the spontaneous-parametric-down-conversion (SPDC) process, which has been the workhorse for quantum photonics sources for many decades [47]. On the other hand, silicon, having a centro-symmetric crystal structure, does not posses a $\chi^{(2)}$. Nevertheless, we can use the $\chi^{(3)}$ non-linearity. Since $\chi^{(3)}$ is a four-indices tensor, it involves mixing between four fields, which means that the general Hamiltonian for third order non-linearities is given by [48, 49]

$$\hat{H}(t) = -\frac{\epsilon_0}{4}\int d\boldsymbol{r}\chi^{(3)}_{ijk\ell}\hat{E}_i(\boldsymbol{r},t)\hat{E}_j(\boldsymbol{r},t)\hat{E}_k(\boldsymbol{r},t)\hat{E}_\ell(\boldsymbol{r},t), \tag{2.4.2}$$

where again summation over all four indices is implied. In a single-mode waveguide, we limit our discussion to the transverse mode supported by the waveguide, propagating along the z direction. In this case, the quantum mechanical description of the total propagating field can be written as [2, 50]

$$\hat{E}(z,t) \propto \int d\omega\ \hat{a}(\omega)e^{i[\beta(\omega)z-\omega t]} + h.c., \tag{2.4.3}$$

where $\beta(\omega)$ is the wavenumber associated to the waveguide mode. Substituting in eq.(2.4.2) we would obtain a general Hamiltonian including 16 terms: all 2^4 possible combinations of creation and annihilation operators arising from the four contributions of the field. Such a Hamiltonian can represent all possible third-order non-linear effects, including third-harmonic generation, cross-phase modulation, etc. The particular effect we are interested in is *spontaneous four-wave mixing* (SFWM), where a bright laser pump field is provided, associated to the first and second fields in eq.(2.4.2). For bright coherent pump fields, we can adopt a semi-classical approximation where we directly substitute the field amplitude α (α^*) instead of the mode operator $\hat{a}$ ($\hat{a}^\dagger$) in eq.(2.4.3). In this scenario, the only non-negligible terms in the Hamiltonian are of the form, $\hat{a}_i^\dagger(\omega_i)\hat{a}_s^\dagger(\omega_s) + h.c.$, giving [51]

$$\hat{H} = -\gamma L\int d\omega_{p_1}d\omega_{p_2}d\omega_i d\omega_s\ \Phi(\omega_{p_1},\omega_{p_2},\omega_i,\omega_2)\alpha_{p_1}(\omega_{p_1})\alpha_{p_2}(\omega_{p_2})\hat{a}_i^\dagger(\omega_i)\hat{a}_s^\dagger(\omega_s) + h.c., \tag{2.4.4}$$

where the constant γ is the *non-linear parameter*, which depends on the susceptibility and on the mode properties [48], and L is the length of the interaction region. This SFWM Hamiltonian represents the conversion of two pump photons, one from each of the two pump fields (labelled p_1 and p_2), into two photons, namely the idler (i) and the signal (s), as shown in Fig. 2.15. Energy conservation implies the condition

$$\omega_{p_1} + \omega_{p_2} = \omega_i + \omega_s, \tag{2.4.5}$$

while momentum conservation results in the presence in eq.(2.4.4) of the *phase-matching* function $\Phi(\omega_{p_1},\omega_{p_2},\omega_i,\omega_2)$, given by

$$\Phi(\omega_{p_1},\omega_{p_2},\omega_i,\omega_s) = \mathrm{sinc}\left(\frac{\Delta\beta L}{2}\right)e^{i\Delta\beta/2}, \tag{2.4.6}$$

where $\Delta\beta = [\beta_{p_1}(\omega_{p_1}) + \beta_{p_2}(\omega_{p_2}) - \beta_i(\omega_i) - \beta_s(\omega_s)]$. The process of SFWM can be performed in two different regimes, depending on how we decide to pump the silicon waveguides. These two approaches are described below.

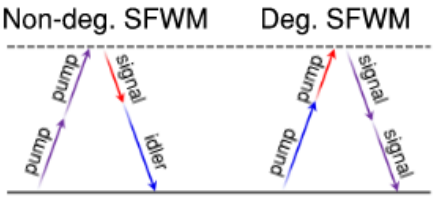


Figure 2.15: Spontaneous four-wave mixing schemes. Left: Non-degenerate SFWM, where two photons are annihilated from the pump, generating a pair of photons with different wavelengths. This effect produces two-mode squeezing. Right: degenerate SFWM, where one photon is annihilated from each of the two pumps with different colours, generating a single pair of degenerate photons. This effect produces single-mode squeezing.

2.4.2 *Generation of squeezing and photon pairs*

In the first regime, called *non-degenerate SFWM*, we use a single input pump, i.e. $p_1 = p_2$. In this case the energy conservation implies that $2\omega_p = \omega_i + \omega_s$, which means that two pump photons from the pump are converted in two photons at different frequencies, symmetric with respect to ω_p (see Fig.2.15). In this case the Hamiltonian can thus be written as

$$\hat{H} = -\gamma L \int d\omega_p d\omega_i d\omega_s \alpha(\omega_p)\alpha(\omega_i + \omega_s - \omega_p)\Phi(\omega_p, \omega_i, \omega_s)\hat{a}_i^\dagger(\omega_i)\hat{a}_s^\dagger(\omega_s) + h.c. \tag{2.4.7}$$

$$= -\gamma LP \int d\omega_i d\omega_s \mathcal{A}(\omega_i, \omega_s)\hat{a}_i^\dagger(\omega_i)\hat{a}_s^\dagger(\omega_s) + h.c. \tag{2.4.8}$$

where $P = |\int d\omega_p \alpha(\omega_p)|^2$ is the total pump power, and the function $\mathcal{A}(\omega_i, \omega_s) = \int d\omega_p \frac{1}{P}\alpha(\omega_p)\alpha(\omega_i + \omega_s - \omega_p)\Phi(\omega_p, \omega_i, \omega_s)$ is the JSA, already introduced in section 2.2.2. Note that all the terms in the integral correspond to two-mode squeezing Hamiltonians, with squeezing parameter given by $|\xi| = \gamma LP$, which linearly depends on the pump power used. Non-degenerate SFWM can, therefore, be used to generate two-mode vacuum squeezing states. The number of independent squeezers acting in the process depends on the factorability of $\mathcal{A}(\omega_i, \omega_s)$ (i.e. on the entanglement in frequency): if it can be factorised in ω_i and ω_s components, for instance with additional filtering, than we can achieve an ideal emission from a single two-mode squeezer.

In the low pumping regime, with squeezing $\gamma LP \ll 1$, the state produced by the non-degenerate SFWM Hamiltonian acting on the vacuum can be approximated with

$$|\psi\rangle \simeq \left(\mathbb{1} - i\gamma LP \int d\omega_i d\omega_s \mathcal{A}(\omega_i, \omega_s)\hat{a}_i^\dagger(\omega_i)\hat{a}_s^\dagger(\omega_s)\right)|0\rangle_i |0\rangle_s \tag{2.4.9}$$

$$= |0\rangle_i |0\rangle_s + |\psi_{11}\rangle_{i,s}\,, \tag{2.4.10}$$

where higher terms in the squeezing expansion are neglected. In this regime, the process is approximating a probabilistic photon-pair source, emitting the bi-photon state

$$|\psi_{11}\rangle_{i,s} = \int d\omega_i d\omega_s \mathcal{A}(\omega_i, \omega_s)\,|1\rangle_i |1\rangle_s \tag{2.4.11}$$

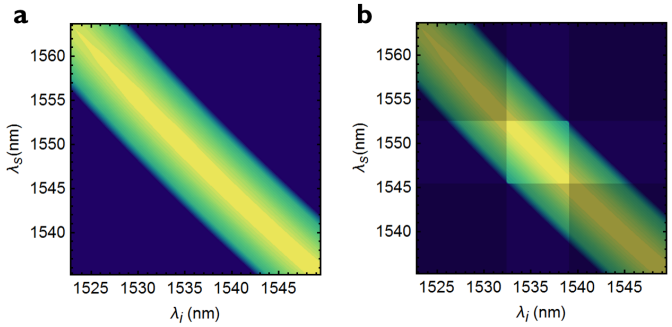


Figure 2.16: Exemplary joint spectra produced in waveguide sources via non-degenerate SFWM. **a,** Unfiltered joint spectrum of the idler and signal photons obtained with a Gaussian pump centred at $\lambda_0 = 1542.9$ nm and bandwidth 1 nm. The purity resulting from this joint spectrum is only 30%. **b,** The purity can be increased by filtering the photons. In this case, rectangular filters of 100 GHz bandwidth are considered, with the shaded areas representing the spectral regions filtered out. The resulting filtered joint spectrum now presents a purity of 96%.

with probability $p = \gamma^2 L^2 P^2$. The sources used in the experiments described in chapters 3, 4 and 6 will be based on this process.

The second regime for SFWM, which will be adopted in the experiment reported in chapter 4, is the *degenerate SFWM* regime (see Fig. 2.15). In this case, a dual pump scheme is used, where pump light is injected at different frequencies, and degenerate photons are emitted. Proceeding analogously as for the non-degenerate SFWM regime, we can obtain that in this scenario single-mode squeezing is obtained, which, in the low pumping scenario, corresponds to the generation of degenerate photon pairs.

In conclusion, to generate squeezed states and single photon states on a chip via SFWM, it is sufficient to propagate the pump through a long silicon waveguide, which intrinsically possesses the required $\chi^{(3)}$ non-linearity. This type of cavity-free sources are generically called *waveguide sources*, and all experiments reported in this thesis are based on them. However, the level of squeezing achievable with this approach is usually quite low (see chapter 4). To increase the efficiency and the purity of the sources, designs based on integrated cavities can be adopted, e.g. via the use of micro-ring resonators [52].

References

1. Kovac, J. M., Leitch, E., Pryke, C, Carlstrom, J., Halverson, N. & Holzapfel, W. Detection of polarization in the cosmic microwave background using DASI. *Nature* **420,** 772 (2002).
2. Gerry, C. & Knight, P. L. *Introductory quantum optics* (Cambridge university press, 2005).
3. Schumaker, B. L. Quantum mechanical pure states with Gaussian wave functions. *Physics Reports* **135,** 317–408 (1986).
4. Olivares, S. Quantum optics in the phase space. *The European Physical Journal* **203,** 3–24 (2012).
5. Weedbrook, C., Pirandola, S., García-Patrón, R., Cerf, N. J., Ralph, T. C., Shapiro, J. H. & Lloyd, S. Gaussian quantum information. *Rev. Mod. Phys.* **84,** 621–669. https://link.aps.org/doi/10.1103/RevModPhys.84.621 (2012).
6. Hanbury-Brown, R & Twiss, R. Q. A test of a new type of stellar interferometer on Sirius. *Nature* **178,** 1046–1048 (1956).
7. Walls, D. F. Squeezed states of light. *Nature* **306,** 141 (1983).
8. Hong, C. K., Ou, Z. Y. & Mandel, L. Measurement of subpicosecond time intervals between two photons by interference. *Phys. Rev. Lett.* **59,** 2044–2046 (1987).
9. Brańczyk, A. M. Hong-ou-mandel interference. *arXiv:1711.00080* (2017).
10. Chen, J, Lee, K. F. & Kumar, P. Deterministic quantum splitter based on time-reversed Hong-Ou-Mandel interference. *Phys. Rev. A* **76,** 031804 (2007).
11. Silverstone, J. W., Bonneau, D., Ohira, K., Suzuki, N., Yoshida, H., Iizuka, N., Ezaki, M., Natarajan, C. M., Tanner, M. G., Hadfield, R. H., *et al.* On-chip quantum interference between silicon photon-pair sources. *Nature Photonics* **8,** 104 (2014).
12. Ulanov, A. E., Fedorov, I. A., Sychev, D., Grangier, P. & Lvovsky, A. Loss-tolerant state engineering for quantum-enhanced metrology via the reverse Hong–Ou–Mandel effect. *Nature Communications* **7,** 11925 (2016).
13. Scheel, S. Permanents in linear optical networks. *arXiv:quant-ph/0406127* (2004).
14. Caianiello, E. R. On quantum field theory—I: explicit solution of Dyson's equation in electrodynamics without use of feynman graphs. *Il Nuovo Cimento* **10,** 1634–1652 (1953).
15. Valiant, L. G. The complexity of computing the permanent. *Theoretical Computer Science* **8,** 189–201 (1979).
16. Aaronson, S. & Arkhipov, A. The computational complexity of linear optics. *Quant. Inf. Comp.* **13,** 1383 (2014).
17. Nielsen, M. A. & Chuang, I. *Quantum computation and quantum information* (Cambridge University Press, 2002).
18. Nolan, J. C. *Synthesis of Multiple Beam Networks for Arbitrary Illuminations* PhD thesis (John Hopkins University, Baltimore, Maryland, 1965).

19. Cummings, W. C. *Multiple Beam Forming Networks* tech. rep. (Lincoln Laboratory, MIT, Boston, 1978).

20. Reck, M., Zeilinger, A., Bernstein, H. J. & Bertani, P. Experimental realization of any discrete unitary operator. *Phys. Rev. Lett.* **73,** 58–61 (1994).

21. Clements, W. R., Humphreys, P. C., Metcalf, B. J., Kolthammer, W. S. & Walmsley, I. A. Optimal design for universal multiport interferometers. *Optica* **3,** 1460–1465 (2016).

22. Russell, N. J., Chakhmakhchyan, L., O'Brien, J. L. & Laing, A. Direct dialling of Haar random unitary matrices. *New Journal of Physics* **19,** 033007 (2017).

23. Pai, S., Bartlett, B., Solgaard, O. & Miller, D. A. Matrix optimization on universal unitary photonic devices. *arXiv:1808.00458* (2018).

24. Carolan, J., Harrold, C., Sparrow, C., Martín-López, E., Russell, N. J., Silverstone, J. W., Shadbolt, P. J., Matsuda, N., Oguma, M., Itoh, M., Marshall, G. D., Thompson, M. G., Matthews, J. C. F., Hashimoto, T., O'Brien, J. L. & Laing, A. Universal linear optics. *Science* **349,** 711–716 (2015).

25. Ralph, T. C., Langford, N. K., Bell, T. B. & White, A. G. Linear optical controlled-NOT gate in the coincidence basis. *Phys. Rev. A* **65,** 062324. https://link.aps.org/doi/10.1103/PhysRevA.65.062324 (2002).

26. Hofmann, H. F. & Takeuchi, S. Quantum phase gate for photonic qubits using only beam splitters and postselection. *Phys. Rev. A* **66,** 024308. https://link.aps.org/doi/10.1103/PhysRevA.66.024308 (2002).

27. Rudolph, T. Why I am optimistic about the silicon-photonic route to quantum computing. *APL Photonics* **2,** 030901 (2017).

28. Knill, E., Laflamme, R. & Milburn, G. J. A scheme for efficient quantum computation with linear optics. *Nature* **409,** 46 (2001).

29. Raussendorf, R. & Briegel, H. J. A one-way quantum computer. *Phys. Rev. Lett.* **86,** 5188 (2001).

30. Raussendorf, R., Browne, D. E. & Briegel, H. J. Measurement-based quantum computation on cluster states. *Phys. Rev. A* **68,** 022312. https://link.aps.org/doi/10.1103/PhysRevA.68.022312 (2003).

31. Briegel, H. J. & Raussendorf, R. Persistent Entanglement in Arrays of Interacting Particles. *Phys. Rev. Lett.* **86,** 910–913. https://link.aps.org/doi/10.1103/PhysRevLett.86.910 (2001).

32. Browne, D. E. & Rudolph, T. Resource-Efficient Linear Optical Quantum Computation. *Phys. Rev. Lett.* **95,** 010501. https://link.aps.org/doi/10.1103/PhysRevLett.95.010501 (2005).

33. Kieling, K., Rudolph, T. & Eisert, J. Percolation, Renormalization, and Quantum Computing with Nondeterministic Gates. *Phys. Rev. Lett.* **99,** 130501. https://link.aps.org/doi/10.1103/PhysRevLett.99.130501 (2007).

34. Gimeno-Segovia, M., Shadbolt, P., Browne, D. E. & Rudolph, T. From Three-Photon Greenberger-Horne-Zeilinger States to Ballistic Universal Quantum Computation. *Phys. Rev. Lett.* **115,** 020502. https://link.aps.org/doi/10.1103/PhysRevLett.115.020502 (2015).

35. Miller, S. E. Integrated Optics: An Introduction. *Bell System Technical Journal* **48,** 2059–2069 (1969).

36. Marcatili, E. A. Dielectric rectangular waveguide and directional coupler for integrated optics. *Bell System Technical Journal* **48,** 2071–2102 (1969).

37. Politi, A., Cryan, M. J., Rarity, J. G., Yu, S. & O'brien, J. L. Silica-on-silicon waveguide quantum circuits. *Science* **320,** 646–649 (2008).

38. Lifante, G. *Integrated Photonics: Fundamentals* (John Wiley & Sons, First edition, 2003).

39. Chang, W. S. *Fundamentals of guided-wave optoelectronic devices* (Cambridge University Press, 2009).

40. Yariv, A. Coupled-mode theory for guided-wave optics. *IEEE J. Quant. Elect.* **9,** 919–933 (1973).

41. Soldano, L. B. & Pennings, E. C. Optical multi-mode interference devices based on self-imaging: principles and applications. *J. Light. Tech.* **13,** 615–627 (1995).

42. Bogaerts, W., Dumon, P., Van Thourhout, D. & Baets, R. Low-loss, low-cross-talk crossings for silicon-on-insulator nanophotonic waveguides. *Optics Letters* **32,** 2801–2803 (2007).

43. Atabaki, A., Hosseini, E. S., Eftekhar, A., Yegnanarayanan, S & Adibi, A. Optimization of metallic microheaters for high-speed reconfigurable silicon photonics. *Optics Express* **18,** 18312–18323 (2010).

44. Ding, Y., Ou, H. & Peucheret, C. Ultrahigh-efficiency apodized grating coupler using fully etched photonic crystals. *Opt. Lett.* **38,** 2732–2734 (2013).

45. Taillaert, D., Bogaerts, W., Bienstman, P., Krauss, T. F., Van Daele, P., Moerman, I., Verstuyft, S., De Mesel, K. & Baets, R. An out-of-plane grating coupler for efficient butt-coupling between compact planar waveguides and single-mode fibers. *IEEE J. Quant. Electr.* **38,** 949–955 (2002).

46. Ding, Y., Peucheret, C., Ou, H. & Yvind, K. Fully etched apodized grating coupler on the SOI platform with -0.58 dB coupling efficiency. *Opt. Lett.* **39,** 5348–5350 (2014).

47. Kwiat, P. G., Mattle, K., Weinfurter, H., Zeilinger, A., Sergienko, A. V. & Shih, Y. New High-Intensity Source of Polarization-Entangled Photon Pairs. *Phys. Rev. Lett.* **75,** 4337–4341. https://link.aps.org/doi/10.1103/PhysRevLett.75.4337 (1995).

48. Boyd, R. W. *Nonlinear optics* (Elsevier, 2nd edition, 2003).

49. Carter, S. J., Drummond, P. D., Reid, M. D. & Shelby, R. M. Squeezing of quantum solitons. *Phys. Rev. Lett.* **58,** 1841–1844. https://link.aps.org/doi/10.1103/PhysRevLett.58.1841 (1987).

50. Blow, K. J., Loudon, R., Phoenix, S. J. D. & Shepherd, T. J. Continuum fields in quantum optics. *Phys. Rev. A* **42,** 4102–4114. https://link.aps.org/doi/10.1103/PhysRevA.42.4102 (1990).

51. Christensen, J. B. *Tailored four-wave-mixing processes for optical quantum information science* PhD thesis (Technical University of Denmark, Lyngby, Denmark, 2018).

52. Silverstone, J. W., Santagati, R., Bonneau, D., Strain, M. J., Sorel, M., O'Brien, J. L. & Thompson, M. G. Qubit entanglement between ring-resonator photon-pair sources on a silicon chip. *Nature Communications* **6,** 7948 (2015).

3 INTEGRATED HIGH-DIMENSIONAL ENTANGLED PHOTONS

We have described in the previous chapter how quantum information tasks can be performed in photonics, and why the integrated approach is a promising platform for future quantum applications. The experiment reported in this chapter is a clear practical demonstration of such capability. Here, we investigate silicon quantum photonics as a platform for large-scale experiments with integrated photon sources and circuits. The focus of the experiment is to develop and control large complex interferometers for the generation and processing of high-dimensional entangled states of photons. A particular result of our experiment is to demonstrate the potential of silicon photonics for developing high-dimensional quantum technologies, i.e. technologies based on qudits.

Multiple experimental approaches have been previously developed for generating and controlling high-dimensional systems. The encoding of qudits has been demonstrated in a wide variety of different photonic degrees of freedom, e.g. orbital angular momentum [2], time-bin [3], and frequency [4]. However, these approaches present some significant challenges: 1) the difficulty in generating and processing high-dimensional states with high-fidelity, and 2) the difficulty in performing arbitrary operations on the qudits. These challenges represent significant limitations, as high-fidelity and universality are important requirements in most quantum applications. In the experiment we will show how an approach based on large-scale integrated quantum optics, amongst other advances, allows us to overcome such challenges.

The chapter starts with a description of the large-scale silicon photonic device and the set-up used in the experiment. We will then report the preliminary characterisation of the device and the reconstruction of the generated high-dimensional entangled states. Finally, implementations of a range of quantum applications based on qudits are presented and discussed.

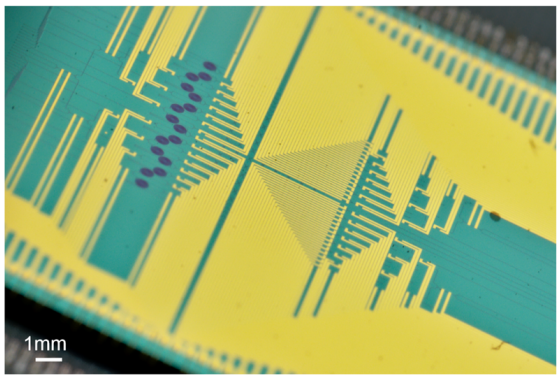


Figure 3.1: Photo of the device. The opaque black lines are the silicon waveguides. The array of 16 SFWM spiral sources can be observed as black dots on the centre-left. The gold wires are the on-chip electrical connections individually accessing the phase-shifters for their control. Wire-bonds (192 in total) that electronically connect the silicon chip to an external PCB can be observed on the sides of the chip.

3.1 Silicon photonic chip: design and characterisation

We start by describing the large-scale silicon photonics device implemented, shown in Fig. 3.1. The general structure of the scheme used to generate and process entangled qudits is pictured in Fig. 3.2, and the actual photonic circuit used is shown in Fig. 3.3. The silicon device designed and fabricated for the experiment represented an important technological step in terms of number of integrated optical components, including photon sources and circuits, coherently utilised on a single photonic chip. Of course, when trying to push such technological limits there are many challenges that need to be addressed.

3.1.1 *Chip details and experimental setup*

As can be observed from the schematic of the circuit in Fig. 3.3, the photonic chip presents a remarkable circuit complexity. In total, more than 500 optical components are integrated on the same device to form a stable large-scale interferometer, comprised of an array of 16 SFWM photon sources, 93 configurable thermal phase shifters, 122 beam-splitters, 136 waveguide crossers and 32 grating couplers. The device is designed and fabricated on commercial silicon-on-insulator (SOI) using electron-beam lithography. The silicon waveguides, including the 1.5 cm long spirally-shaped waveguide sources, have a cross-section of 450 nm $\times$ 260 nm. The beam-splitters were based on multi-mode

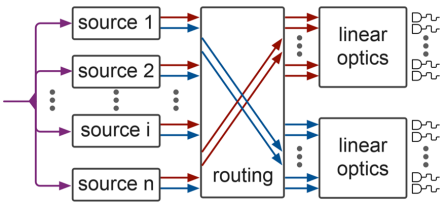


Figure 3.2: Representation of the scheme for the generation and manipulation of bipartite states with tunable dimension and entanglement.

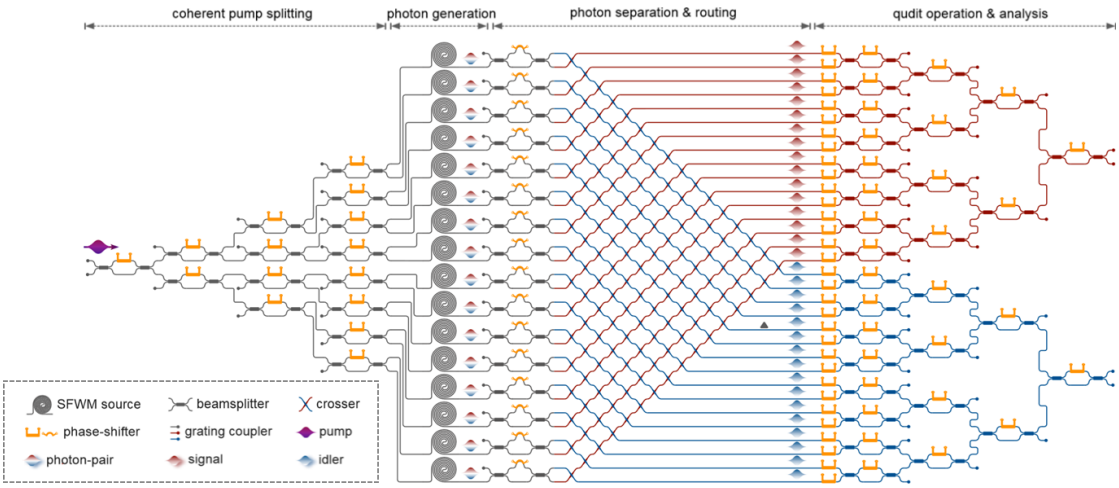


Figure 3.3: Photonic circuit diagram. The input pump light is initially split by a network of MZIs between an array of 16 SFWM waveguide sources. A pair of photons pair is produced in superposition between the 16 sources, generating a reconfigurable high-dimensional bipartite entangled state. The two generated photons with different wavelengths are separated by an array of AMZI filters and routed by a structure of waveguide crossers. Arbitrary local projective measurements are then performed via triangular networks of MZIs.

interferometers (MMIs), described in section 2.3.1, and combined with phase shifters, based on titanium heaters, to form the MZIs. These components are also used to form AMZI filters to separate idler and signal photons. The insertion loss for the MZIs was measured to be below -0.2 dB, and typical extinction ratio in the interference filers in the MZIs was approximately 30 dB, indicating good 50:50 balance of the MMI splitters. The waveguide crossers are designed using the self-imaging effect in multi-mode interferometers, and presented insertion loss lower than -0.1 dB with crosstalk well below -40 dB. The coupling of light into and out of the chip is performed via grating couplers based on Bragg interference in a photonic crystal. At the wavelengths used (near-1550 nm) the fabricated grating coupler presented a coupling efficiency of -2.4 dB. The propagation loss in the waveguides was estimated to be -3.1 dB/ cm, which was, however, significantly higher in the spirally-shaped waveguides forming the sources, mostly due to bending losses in the spirals and difficulties in fabricating circular-shaped elements with electron-beam lithography (rectangular-shaped spirals should be preferred, at the cost of slightly worse miniaturisation). Total insertion losses of the device were measured to be approximately -29 dB.

In the experiment two photons are used, where, as described in more details below, the

photon pair is generated in a superposition between the 16 spiral sources. In particular, we injected into the chip CW pump light with wavelength of $\lambda_p = 1544.49$ nm to generate photon pairs, signal and idler, in a broadband spectral distribution (≈ 30 nm bandwidth) via non-degenerate SFWM. We post-selected the signal photon at $\lambda_s = 1539.73$ nm and idler photon at $\lambda_i = 1549.32$ nm, by combing the on-chip AMZI filters and off-chip filters. The latter also presented high extinction (> 100 dB) to additionally remove spurious pump light. The photons, once collected out of the chip via a fibre array, are finally detected off-chip via two superconductive nanowire singe-photon detectors (SNSPDs) with average efficiency of 85%. In standard measurement conditions, where ≈ 40 mW input CW pump power was injected, the count rate measured at the detector was approximately 2 kHZ, with negligible accidental counts (signal-to-noise ratio > 100).

A significant challenge in developing this large-scale silicon quantum photonic experiment was the active control of the device, that is the simultaneous and individual electronic driving of the 93 phase shifters. To achieve that, the electronic pads from the chip were wire-bonded to a PCB and connected to a computer-interfaced digital electronic controller with 96 channels and 12-bit resolution on the voltage. A Mathematica interface was developed to actively configure the device by controlling both the phase shifters and the data collection. This allowed to configure the whole chip, collect and process the data, and decide the next configuration for the chip at kHz rate, limited by the thermo-optical response time of the phase shifters. This rate enabled us to perform long data collections for tomographic reconstruction of the generated states, as described in the next sections.

For each heater a voltage of approximately 5 V was required to achieve a 2π phase shift, corresponding to ≈ 50 mW of thermal power dissipated by the heater. In total, for configurations which required all 93 phase-shifters to be set, typically around 1 W of power was dissipated on the chip. To maintain a constant temperature of the chip, e.g. to avoid de-coupling due to a thermal relaxation of the device, a Peltier-cell together with a thermistor and a proportional integrative derivative controller were used. A standard water-cooling system was built to ensure an efficient dissipation of the thermal power extracted by the Peltier-cell.

3.1.2 *Generation of photonic high-dimensional entanglement*

As we have just seen, current semiconductor technologies allow us to integrate hundreds of optical components on a single large-scale photonic device to form complex networks of photon sources and interferometers. However, great inventiveness is required to develop linear-optical schemes where such complex interferometers actually perform the desired task. For this experiment a scheme was devised for the generation and manipulation of high-dimensional quantum states. The scheme, reported in Fig. 3.2, is quite general, and can be used to generate and operate bipartite states in any dimension and with an

arbitrary degree of entanglement.

For the production of entangled qudits, the idea is to generalise the approach for generating entangled path-encoded qubits on-chip, which is based on the coherent pumping of two integrated sources [5–10]. In fact, in a similar fashion, the generation of two entangled d-dimensional qudits can be straightforwardly obtained by pumping d sources [2, 11]. This idea, pictured in Fig. 3.2, was originally investigated in Ref. [11] in the context of path-encoded qudits in fibre-based photonic architectures. However, it required an optical stability which is challenging for optical fibres, and limited the implementation to bipartite two-dimensional systems, i.e. standard qubits. On the other hand, it represents a natural approach for integrated quantum photonics, where the stability is intrinsic and large interferometers and array of sources can be deployed. The approach can thus be readily adapted to integrated circuits as follows.

As pictured in Fig. 3.3, we inject the pump light into the chip and coherently split it across the array of 16 SFWM sources via a network of cascaded MZIs. The MZIs can be reconfigured through the phase shifters, so that the pumping power arriving at each source can be controlled. As the pump is a CW laser, the sources are operated in the low-squeezing regime, so that to a very good approximation we can focus on the case where a single pair of photons is produced. Crucially, the pair of signal-idler photons (with different frequencies) must originate from the same source, so that the state generated is given by $\sum_{k=0}^{d-1} c_k \left|1\right\rangle_{i,k} \left|1\right\rangle_{s,k}$ where $\left|1\right\rangle_{i,k}$ ($\left|1\right\rangle_{s,k}$) indicates the Fock state of the idler (signal) photon being in its k-th spatial mode. The coefficients c_k are the complex amplitudes associated to the different modes (with $\sum |c_k|^2 = 1$ for normalisation). At this point, we can simply map the Fock state of the photons into a logical high-dimensional system using the encoding presented in section 2.2.5. That is, we label the Fock state associated to a photon in its k-th optical mode as the logical qudit state $\left|k\right\rangle$ ($k = 0, \ldots, d-1$). In this way we obtain the high-dimensional bipartite entangled state

$$\left|\psi_d\right\rangle = \sum_{k=0}^{d-1} c_k \left|k\right\rangle_i \left|k\right\rangle_s . \tag{3.1.1}$$

Note that, essentially, what is happening is that we are converting, through the SFWM process, the classical coherence of the pump laser light across the sources into a quantum coherence in the superposition of the emitted photons. The amplitudes c_k can be configured in arbitrarily by controlling the pump splitting over the d sources and the individual phases on each mode, which is performed by reconfiguring the MZIs at the input. In particular, $|c_k|^2$ is proportional to the square of the pump power at the k-th source (see section 2.4). This allows us to tune the level of entanglement generated. For example, in the case where the pump is uniformly split across the sources (e.g. by setting

all the input MZIs as 50:50 beam-splitters) all coefficients are constant, yielding the maximally entangled state of two qudits

$$\left|\psi_d^+\right\rangle = \frac{1}{\sqrt{d}} \sum_{k=0}^{d-1} |k\rangle_i |k\rangle_s . \tag{3.1.2}$$

In contrast, if we direct all the pump into a single source, for instance the k-th, we obtain the separable state $|k\rangle_i |k\rangle_s$ which doesn't present any entanglement. Similarly, note also that we can decide how many sources to pump, tuning the dimensionality of the system; in our device, this goes up to a maximum number of $d = 16$ sources. After the generation, the array of AMZIs filters is used to deterministically divide the idler and signal photons, and a network of waveguide crossers allow us to separate the modes associated to the different photons. At this point, the two qudits, with the associated modes grouped together, can be locally addressed.

As a final remark on the generation scheme, it is interesting to investigate how the generation rate changes for different dimensions. As described in section 2.4, for SFWM sources in the low-power pumping regime the probability to emit a photon pair is proportional to the square of the pump power. For a fixed amount of power coupled into the chip, when splitting the pump through d sources (for simplicity we consider a uniform distribution of the pump) the power at each source decreases as $1/d$, so that for each source the probability to emit a photon pair decreases as $1/d^2$. However, we now have a combinatorial enhancement given from the fact that a qudit is generated whenever one over the d sources fire, which can happen in d different ways (such combinatorial enhancement will be generalised to scattershot-type sources in chapter 4). Therefore, for a fixed amount of input pump power, the generation rate for the bipartite entangled states will overall decrease as $1/d$ using the scheme as above.

3.1.3 *Scheme for qudit local manipulation*

The manipulation of high-dimensional states is a notably impractical task in most experimental platforms. For example, for qudits encoded in the OAM degree of freedom of the photons, schemes to perform universal operations are known only for low dimensionalities [12, 13], and were found very recently via learning algorithms [14]. Generalising and implementing such schemes to high dimensions is currently an open problem for the OAM approach. In contrast, as described in section 2.2.4, integrated optics is the only platform where universal linear optical schemes able to perform arbitrary operations in high dimensions have been demonstrated (currently up to $d = 6$) [15]. As pictured in the schematic of Fig. 3.2, such universal schemes could be used in our chip to perform arbitrary local operations on the qudits, which, together with the ability to tune the

degree of entanglement, would allow to generate any bipartite state for $d \leq 16$. However, in our implementation, to reduce the number of elements required, we used a simplified scheme which can perform arbitrary projective measurements. Of course the ability to perform universal measurements is less stringent than performing universal operations. In particular, as described below, arbitrary measurements on d-dimensions require $\mathcal{O}(d)$ elements, in contrast to $\mathcal{O}(d^2)$ for universal operations. For instance, this allowed us to perform the experiment with "only" 93 total reconfigurable phase-shifters, instead of the 511 we would have required to perform the experiment with universal unitary manipulations.

The scheme to perform arbitrary local projective measurements comprises an array of phase-shifters followed by a triangular network of MZIs, as shown in Fig. 3.3. Projective measurements on the computational basis of a path-encoded photonic qudit are simply obtained by detecting photons in the optical spatial mode associated with the state $|k_0\rangle$, and can thus be measured setting all the MZIs as identities. A projection onto an arbitrary state $|\tilde{\psi}\rangle = \sum_{k=0}^{d-1} a_k |k\rangle$, with $\sum |\alpha|^2 = 1$, is obtained by an operation $|\tilde{\psi}\rangle \to |k\rangle$ followed by a measurement in the computational basis, which can be performed in the triangular structure of MZIs using the following algorithm. The idea is to represent the target state $|\tilde{\psi}\rangle$ as a vector of its computational basis coefficients $\boldsymbol{\alpha}$, and to iterate vector elements elimination in the following way. Up to adding ancillary optical modes, without loss of generality we can consider d to be a power of 2, i.e. $d = 2^N$ for some $N \geq 1$. In this case, the triangular structure will consist of N layers of MZIs, with the ℓ-th layer consisting of $2^{N-\ell}$ MZIs in parallel (where $\ell = 1, \ldots, N$). The values of the phases inside the MZIs and in the array of phase-shifters on each mode before the structure are chosen such that, after the ℓ-th layer, all the elements of the evolved vector $\boldsymbol{\alpha} \to \boldsymbol{\alpha}'^{(\ell)}$ with index k, such that $\lfloor k/2^{(\ell-1)} \rfloor \mod 2 \equiv 1$, are set to zero. This elimination procedure, similar to the one performed for universal decompositions, is possible by iterating MZI operations and via simple trigonometric calculations. Taking $k_0 = 2^{N-1}$ as the output mode, the only element which can be non-zero after the N elimination steps is $\alpha'_{k_0} = 1$ associated with the optical mode k_0, and therefore the transformation $|\tilde{\psi}\rangle \to |k_0\rangle$ is obtained. Note that this operation is deterministic and no post-selection is required. In our experiment, we implement the correlation measurements by collapsing the states into the $|k_0 = 7\rangle$ computational basis state.

The number of elements required to implement the scheme on a system of size d can be readily calculated. As before, we can consider for simplicity $d = 2^N$ for some $N = \log_2 d \geq 1$. The scheme then comprises d initial phase-shifters and N layers of MZIs, with the ℓ-th layer consisting of $2^{N-\ell}$ MZIs. The total number of elements is then given by

$$d + \sum_{\ell=1}^{N} 2^{N-\ell} = d\left(1 + \sum_{\ell=1}^{\log_2 d} 2^{-\ell}\right) = d\left(2 - \frac{1}{d}\right) = 2d - 1 = \mathcal{O}(d). \tag{3.1.3}$$

Note that similar calculations can also be performed for the generation scheme, which also need $\mathcal{O}(d)$ components. The approach is thus general for arbitrary dimensions, with the complexity of the integrated devices that scales linearly with the dimension of each quantum system (quadratically when using universal operations).

3.1.4 *Technical challenges in the experiment*

As in any experiment, while operating the device, we encountered a variety of challenges that needed to be addressed, mostly accountable to the hardness of controlling such complex device. They can be grouped in three main classes.

The first concerns the fabricability of large-scale devices of similar size, or larger, to the one reported here. The fact that current technologies allow to manufacture photonic chips of such complexity is already an achievement. However, the limits of current fabrication capabilities played a role already in our experiment. For example, a first version of the device, fabricated approximately two years before the publication of the experiment (spring 2016), was unusable. The main problem with it was excessive photon loss in the chip. This issue was mostly due to fabrication imperfections when fabricating the spiral sources using electron-beam lithography, resulting in total insertion losses higher than -40 dB. This implied count rates of approximately 1 event/minute when pumping 4 sources in the array, making the experiment unfeasible. We then designed and fabricated a second version of the device, which is the one we used for the final experiment. Here the losses were significantly lower and the count rates enough to perform the experiment. However, a fabrication error still occurred: due to a dust particle accidentally deposited during the waveguide fabrication, one of the modes in the network of crossers presented -10 dB additional losses (pictured with a black triangle in Fig. 3.3), making one of the 16 available dimensions unusable. In practise, this limited us to study entangled systems up to dimension $d = 15$.

The second class of problems came from having to precisely control a large number of reconfigurable heater-based phase shifters on a mm-scale device. As described in section 3.1.1, the electronic control of these was achieved via an external computer-controlled voltage driver. However, the thermo-optic response of each single heater had to be precisely characterised. Given the complexity of the device, and the difficulty of isolating single elements in integrated circuits, devising a characterisation routine where each phase shifter was characterised independently from the others was quite a non-trivial task. The characterisation procedure we designed, would typically take around 10 hours,

and had to be repeated approximately once per week to correct drifts in the heaters and maintain a good control of the device. A large part of the time spent on this experiment was thus dedicated to the characterisation of the heaters. As this issue will become more significant increasing the system size, for future large-scale experiments it will be important to design the chip explicitly so that an automatic characterisation procedure is supported and can be efficiently executed.

Thermal cross-talk between the heaters represents a significant source of error in integrated operations. As the chip size in our case was relatively large (~ 2 cm^2) compared to standard silicon photonics chips, thermal cross-talk was smaller than what is typically observed in silicon devices. We also mitigated it via maintaining a stable temperature on the chip via the cooling system described in 3.1.1. As characterised in the next sections, this allowed us to maintain a precise control of the phases.

The third type of issues encountered is instead due to indistinguishability of the photons emitted from different sources. In particular, source distinguishability would introduce noise in the generated states in eq.(3.1.1), decreasing their purity. While indistinguishability due to temporal mismatch doesn't play a role in integrated circuits (path differences are compensated in the circuit design), spectral mismatches still represent a problem. We mitigated this issue by using off-chip narrow filters (1 nm bandwidth) for the photons. The indistinguishability was tested performing reverse Hong-Ou-Mandel interference between all possible pairwise choices of the sources, as shown in the next section.

3.1.5 *Characterisation of the sources and measurements*

The characterisation of the phase-shifters, operations and sources is a central challenge for large-scale integrated optical experiments. For each thermo-optical phase-shifter, the characterisation can be obtained in two steps:

1. Characterisation of the electrical response, measuring the current-voltage curve for each heater. A parabolic function was used to fit the I-V curves to include a non-ohmic response.

2. Optical characterisation, mapping electrical power to optical phase. This is obtained by measuring the output optical power as a function of the electrical power dissipated in each phase-shifter. The measured classical fringe is fitted to a sinusoid function.

By repeating these two steps, each phase-shifter can be fully characterised, determining the associated optical phase as a function of the applied voltage. Figure 3.4a shows one example of classical interference used for the heater characterisation. To characterise local phase-shifters that are not directly included in an interferometer, e.g. the ones

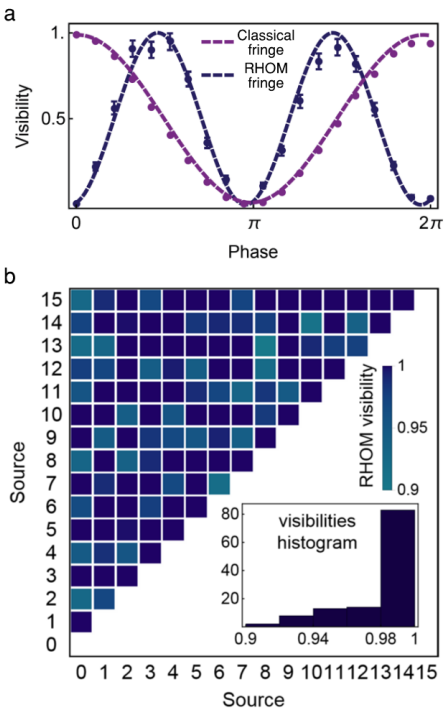


Figure 3.4: Classical and quantum fringes. **a**, Example of classical interference (purple) and two-photon quantum interference (blue), obtained via the rev-HOM effect, measured when scanning a phase shifter in the output MZIs. The points are experimental data, while the lines are sinusoidal fittings. Error bars are given by photon Poissonian statistics. In the classical case, the fringes have been performed for all 93 heaters to characterise their thermo-optic response. In the quantum case, the rev-HOM fringes have been performed for all 120 possible pairs of the 16 sources to test their indistinguishability, with visibilities shown in **b**. The 120 experimental visibilities between all possible combinations of sources, with an average visibility of 0.984 ± 0.025, are shown in the inset.

before the triangular network of MZIs, the MZIs in the input and output networks need to be reconfigured so that the desired phase-shifter becomes embedded inside a large MZI. The optical phase applied can then be determined from the interference fringe associated to the embedding MZI.

Similarly, we can configure the input and output networks of MZIs to embed two chosen sources on a large MZIs. In this way, by scanning the local phase associated to the mode of one of the two sources, and performing single-photon detection at the output, we can obtain a quantum interference fringe via the rev-HOM effect described in section 2.2.2. An example of measured rev-HOM fringe is reported in Fig. 3.4a together with the classical fringe obtained from the same interferometer (note the doubling of the fringe frequency in the quantum interference case). By performing rev-HOM interference between all $\binom{16}{2} = 120$ possible pairs of sources, where each pair can be selected by simply reconfiguring the input and output MZI networks, we obtain a pairwise characterisation of the sources indistinguishability. In particular, as discussed in section 2.2.2, the visibility represents the overlap between the spectra of the different sources used. The measured visibilities of the rev-HOM fringes are reported in Fig. 3.4b. The experimental achieved visibilities were >0.90 in all cases, with more than 80% cases having a visibility >0.98.

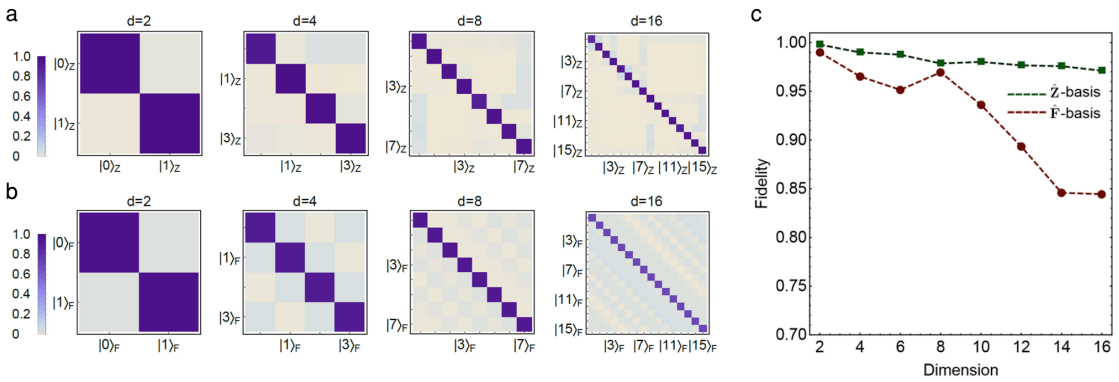


Figure 3.5: Preliminary characterisation of high-dimensional operations. **a**, Measured probabilities distributions preparing and projecting the states on the computational $\hat{Z}$-basis for qudits with dimension 2, 4, 8 and 16. The horizontal axes refer to the prepared input states and the vertical axes to the measured output ones. **b**, The input/output probabilities are shown for states prepared and measured in the Fourier $\hat{F}$-basis. **c**, Mean statistical fidelities for d-dimensional projectors, in both the computational $\hat{Z}$-basis (green) and the Fourier $\hat{F}$-basis (red).

These results show a good degree of indistinguishability between the sources.

The characterisation of the phase shifters was done in a regime where each of them was singularly addressed. However, the main source of noise in the operation, i.e. thermal cross-talk, manifests itself in the collective behaviour of the heaters. In order to characterise this type of noise we studied the fidelities of the input-output probability distributions for the qudit projectors, where multiple heaters have to be simultaneously controlled. We tested the projection fidelity in dimension $d = 2$ to 16 for the computational basis $\hat{Z} = |k\rangle\langle k|$ and the Fourier-transform basis $\hat{F} = |\ell\rangle\langle\ell|$. Here $|\ell\rangle = \sum_{k=0}^{d-1} e^{2\pi i k\ell/d}|k\rangle/\sqrt{d}$ and $k, \ell = 0, \ldots, d-1$. Figure 3.5a shows the measured distributions in the $\hat{Z}$-basis, while those for the $\hat{F}$-basis are reported in Fig. 3.5b.

Figure 3.5c shows the classical mean statistical fidelities, defined as $\sum_i \sqrt{p_i q_i}$ where p_i and q_i are the measured and ideal output probabilities, for different dimensions. We obtain fidelities 0.998, 0.990, 0.979 and 0.971 for $d = 2$, 4, 8 and 16 on the $\hat{Z}$-basis and fidelities of 0.990, 0.965, 0.970 and 0.844 for $d = 2$, 4, 8 and 16 on the $\hat{F}$-basis. Note that the results for the $\hat{F}$-basis show higher noise in the qudit operations compared to the $\hat{Z}$-basis. The main cause of this is residual thermal cross-talks between the heaters. In fact, measurements on the $\hat{F}$-basis required more heaters to be turned on and rely significantly on the phases before the output triangular MZI networks, for which we expect higher disturbance (the distance between the heaters is smaller). Measurements in the $\hat{F}$-basis are thus more susceptible to cross-talk errors, which is reflected in lower mean statistical fidelities.

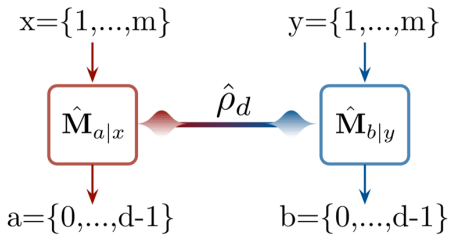


Figure 3.6: Framework for correlation measurements on a bipartite d-dimensional state $\hat{\rho}_d$. The operators $\hat{M}_{a|x}$ and $\hat{M}_{b|y}$ indicate local projective measurements x and y on Alice and Bob, respectively, with associated outcomes a and b. In the experiment reported here, Alice is represented by the signal photon, and Bob by the idler.

3.2 Results

Having developed a precise control of the device, we achieved the ability to generate states with a tunable dimension and degree of entanglement, and to perform arbitrary local measurements which could be rapidly and precisely reconfigured. The most enjoyable part of the experiment began: we had plenty of ideas and applications to implement, and the device readily allowed us to investigate them with a good precision. In general, we focused on applications that could both display the technical advancements and the potential of integrated quantum photonics as a platform for high-dimensional quantum technologies. Many of the applications demonstrated required an advanced theoretical support. For the scope of this thesis, we will focus here on the experimental results, postponing the main theoretical details in appendix C.

3.2.1 *Dimensionality certification: dimension witnesses*

The dimension of a quantum system represents its ability to encode information and constitutes an important resource for quantum information processing. In the previous sections, we have claimed that we can generate states with local dimensions up to $d = 15$, obtained by pumping 15 physical sources. However, an important question both from a fundamental and an application point of view is how to rigorously provide evidence that the dimension in our real system is in fact at least d. Dimension witnesses have been developed to address this point. In particular, device-independent (DI) dimension witnesses enable us to pose an experimental lower bound on the underlying dimension of a quantum system using only the measured correlation probabilities $p(ab|xy)$ [16, 17].

The context for such correlations probabilities in our implementation is depicted in Fig. 3.6, which shows the experiment in the framework of bipartite correlation measurements. The correlations between two parties Alice (A) and Bob (B), here identified by the signal and idler photons, respectively, are represented by the joint probabilities $p(ab|xy) = \mathrm{Tr}[\hat{\rho}_d(\hat{M}_{a|x} \otimes \hat{M}_{b|y})]$, where $\hat{\rho}_d$ is the bipartite state with local dimension d, $x, y \in \{1, \ldots, m\}$ are the m measurement settings Alice and Bob can chose, and $a, b \in \{0, \ldots, d-1\}$ label the possible outcomes. The operators $\hat{M}_{a|x}$ and $\hat{M}_{b|y}$ represent

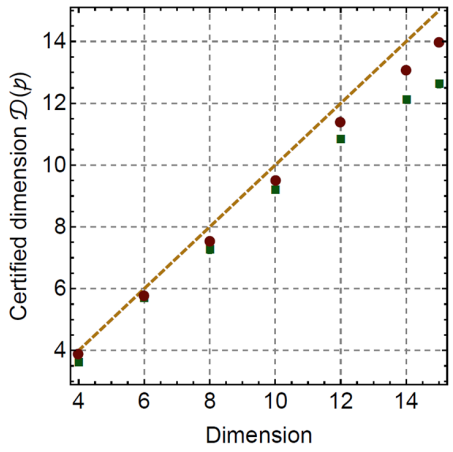


Figure 3.7: Lower bound on the system dimension via measured dimension witnesses. Data points refer to the measured lower bounds on the local dimension of the generated entangled states; red (green) points represent data for the measurement scenario I (II). The yellow line refers to ideal values. Error bars, obtained propagating Poissonian photon statistics, are smaller than the markers and neglected in the plot for clarity.

the projections associated to these measurements.

For the dimension witness, we adopted the approach of Ref. [17], which allows us to extract a lower bound on the local dimension of entangled states from an arbitrary set of measured correlations $p(ab|xy)$. The lower bound takes the form of $d \geq \lceil \mathcal{D}(p) \rceil$, where $\lceil \epsilon \rceil$ indicates the least integer $\geq \epsilon$, and $\mathcal{D}(p)$ is a non-linear function of the correlations given by

$$\mathcal{D}(p) = \left[\sum_{b,b'} \left(\sum_{a} \sqrt{p(ab|xy)p(ab'|xy')} \right)^2 \right]^{-1} . \tag{3.2.1}$$

As the lower bound can be calculated for any set of measured correlations, the experimenter has a freedom in the choice of the measurement bases x, y and y'. In particular, we adopted two different measurement scenarios. In the first scenario, we prepare the d-dimensional maximally entangled state $|\psi_d^+\rangle$ for $4 \leq d \leq 15$ and measure each of the two local systems in the $\hat{Z}$-basis ($x = y = y' = \hat{Z}$). In the second scenario, we tested correlations arising from quantum games [18], i.e. preparing again the maximally entangled state $|\psi_d^+\rangle$ for $4 \leq d \leq 15$ and performing the measurements corresponding to optimal strategies for Magic Square and Magic Pentagram games (see e.g. [18, 19]). Such measurements, described more in details in appendix C, involved computational basis measurements for x and y, while y' required pairwise coupling between the output modes.

Experimental results for these dimension witnesses are shown in Fig. 3.7. Note that in the first scenario we are able to obtain a higher bound, which is due to lower noise in this measurement settings (computational basis measurements required to control fewer heaters, see section 3.1.1). In particular, in the first case, the experimentally observed correlations p_d^{I} yield $\lceil \mathcal{D}(p_d^{\mathrm{I}}) \rceil = d$ for all $d \leq 14$, certifying the ideal dimensions. This

allows to claim that we can generate systems up to dimension 14 (does not certify that they are entangled though!). Note that the fact that for $d = 15$ we obtain a lower bound of 14 does not mean that the system, obtained by physically pumping 15 sources and controlling 15 local modes, does not really have local dimension $d = 15$. Rather it indicates that, due to imperfect correlations, the information encoded in the global state could have been compressed into a bipartite system with local dimension $d = 14$.

3.2.2 *Entanglement certification: violation of high-dimensional Bell and EPR steering inequalities*

The main claim of the experiment is to generate and manipulate high-dimensional entanglement. In the previous section we have demonstrated that the bipartite systems generated are in fact high-dimensional, so we are now left to validate that the system is entangled. As entanglement witnesses we use the violation of high-dimensional inequalities related to two types of quantum correlations: Bell and EPR steering inequalities. While Bell inequalities test non-local correlations between different parties, i.e. correlations that cannot be described by classical local-hidden variables (LHV) theories [20], EPR steering refers to the possibility of one observer to change the state of another distant party by applying local measurements [21, 22]. In particular, EPR steering correlations are incompatible with a classical-quantum model where quantum states, locally predetermined, are sent to different parties, known as local-hidden state (LHS) models. In both cases, the violation of an inequality is used to demonstrate the presence of the associated correlations. Furthermore, a hierarchy of such correlations exist: Bell non-locality of a quantum state always implies EPR steerability, which in turn implies the presence of entanglement (non-locality $\Rightarrow$ EPR steering $\Rightarrow$ Entanglement), while the opposite implications are not in general true (non-locality $\not\Leftarrow$ EPR steering $\not\Leftarrow$ Entanglement) [21].

The presence of entanglement can thus be certified by violating Bell and EPR steering

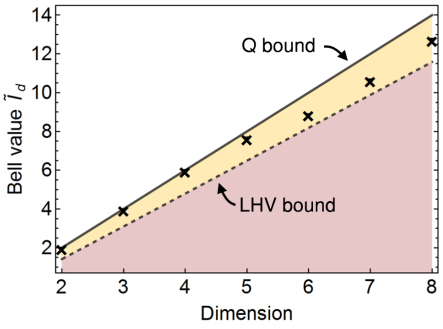


Figure 3.8: Experimental test of the SATWAP Bell-type inequalities for high-dimensional states. Black crosses represent the measured $\tilde{I}_d$ values. Non-locality is certified by violating $\tilde{I}_d \leq C_d$. The dashed line represents the classical LHV bound C_d, while the solid line represents the maximal quantum violation (Tsirelson bound) Q_d. Errors, calculated assuming photon Poissonian statistics, are smaller than markers.

inequalities: they work as entanglement witnesses. From an application point of view, the violation of Bell inequalities is referred to as a device-independent entanglement certification, i.e. the entangled state source and both parties are all considered untrusted, while a violation of EPR steering inequalities as a one-sided device-independent entanglement certification, i.e. one of the two parties is now trusted. Clearly, a device-independent certification of entanglement is more stringent then the one-sided counterpart, and consequently more experimentally demanding in terms of noise tolerance.

For the certification of Bell non-locality in our bipartite high-dimensional system we studied a novel type of multidimensional Bell inequalities recently proposed [23], which we named SATWAP after the authors (Salavrakos-Augusiak-Tura-Wittek-Acin-Pironio). In contrast to previous standard Bell inequalities for high-dimensional systems [24], the SATWAP inequalities have the important feature of being maximally violated by maximally entangled bipartite states of qudits in any dimension. This increases the robustness to experimental noises, and, as we will shortly see, enables interesting applications [23]. The SATWAP inequalities for a bipartite system of local dimension d are of the form $\tilde{I}_d \leq C_d$, where

$$\tilde{I}_d = \sum_{i=1}^{2} \sum_{l=1}^{d-1} \langle A_i^l \bar{B}_i^l \rangle, \tag{3.2.2}$$

and $C_d = [3\cot(\pi/4d) - \cot(3\pi/4d)]/2 - 2$ is the classical bound. The $2(d-1)$ terms in the sum $\langle A_i^l \bar{B}_i^l \rangle$ are generalised Bell correlators. Their calculation requires measurements in the Fourier bases for Alice and Bob. The maximum value of $\tilde{I}$ achievable with quantum states Q_d (namely the Tsirelson bound), can be calculated for arbitrary dimensions $\tilde{I}_d \leq Q_d = 2d - 2$ and is obtained with maximally entangled states [23]. More details for the inequalities and the explicit form of the generalised correlators can be found in appendix C.

The experimental values of $\tilde{I}_d$ are shown in Figure 3.8 for dimensions 2 to 8, together with the analytical maximal bounds for the quantum and classical case. In all tests we

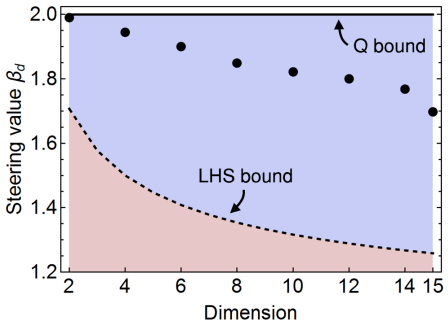


Figure 3.9: The presence of high-dimensional EPR steering correlations is verified via the violation of the inequality $\beta_d \leq \beta_{lhs}$ The dashed line represents the classical bound for LHS models β_{lhs}. Black points are experimental values for β_d. Error bars, calculated assuming Poissonian photon statistics, are smaller than markers.

observe a violation of the classical bound. In particular, a strong violation is achieved for dimensions 2–4, closely approaching the quantum bound Q_d. This certifies, in a device-independent manner, that all the generated states are entangled up to $d = 8$.

For EPR steering correlations, we use EPR steering inequalities of the form $\beta_d \leq \gamma_C$, where

$$\beta_d = \sum_{\substack{a=b \\ x=y}} p(a|x)\mathrm{Tr}[\hat{M}_{b|y}\hat{\rho}_{a|x}]. \tag{3.2.3}$$

and $\gamma_C = 1+1/\sqrt{d}$ is the classical bound. Here $p(a|x)$ represents the probabilities of Alice's measurements; $\hat{M}_{k|0} = |k\rangle\langle k|$, $\hat{M}_{\ell|1} = |-\ell\rangle\langle -\ell|$ are the characterised measurements of Bob, with $|k\rangle$ corresponding to the computational $\hat{Z}$-basis and $|\ell\rangle$ to the Fourier $\hat{F}$-basis, described in previous sections; the state $\hat{\rho}_{a|x}$ represent the reduced state for Bob when the measurement x is performed on Alice and outcome a is obtained. For these inequalities, the maximal quantum violation is 2, constant for all dimensions, and obtained via maximally entangled states [22].

Experimental results for EPR correlations are shown in Fig. 3.9, where measured values of β_d are reported up to dimension 15, all violating the LHS classical bound. This certifies, in a one-side device-independent manner, that all the generated states are entangled up to $d = 15$.

Note that, although the protocols we demonstrate are device-independent or semi-device-independent, in our implementation we actually relied on additional assumptions that are not in general included in device-independent scenarios. In particular, loopholes, such as the detection and locality ones, were not closed in our implementations. Therefore, the protocols reported here and in the following sections cannot be considered as complete realisations of device-independent protocols. However, the main scope of using these protocols here is to employ them as diagnostic tools to characterise our device. As such, in our scenario, the extremely strict requirements for practical field deployment of device-independent protocols (where one has to deal with true Eavesdroppers attacking the experiment) are arguably less relevant.

Up to these minor caveats, we now have rigorous supporting evidence (via entanglement and dimension witnesses) to be able to claim that we have a bipartite entangled quantum system in high-dimensions.

3.2.3 *Quantum state tomographies via compressed sensing*

The most general characterisation of the generated states is the experimental reconstruction of the density matrix, which can be achieved via quantum state tomography (QST). While QST represents an important diagnostic tool to characterise quantum systems, the practical full implementation of QST becomes extremely challenging already

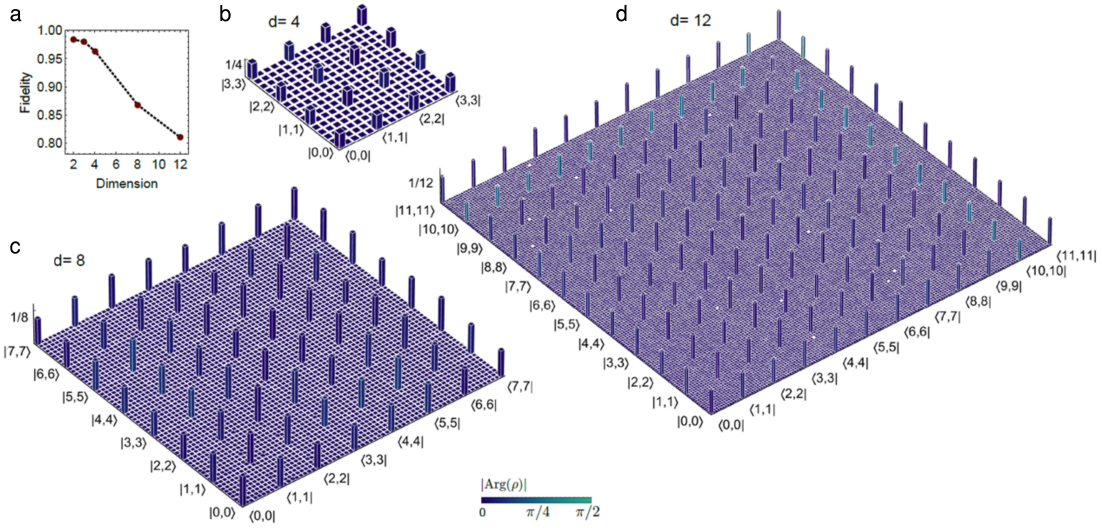


Figure 3.10: State tomographies. **a**, Measured quantum fidelities $\langle\psi_d^+|\hat{\rho}_d|\psi_d^+\rangle$ of the density matrix $\hat{\rho}$ reconstructed via compressed sensing quantum state tomography for different dimensions, where $|\psi_d^+\rangle$ is the ideal bipartite maximally entangled state of local dimension d. Figures **b**, **c** and **d** Represent the experimental density matrix for the d-dimensional entangled state with $d = 4$, $d = 8$ and $d = 12$, respectively. Column heights represent absolute values of the matrix entries $|\rho_{ij}|$, while colours represent the phases $|\arg(\rho_{ij})|$. The phase information for matrix elements with module $|\rho_{ij}| < 0.01$ is approximately randomly distributed and not displayed for more clarity.

for intermediate scale systems, both in terms of the number of measurements required (for a bipartite system of local dimension d it scales as d^4) and the computational time needed to reconstruct the density matrix from the data. In order to perform the tomographic reconstructions of the generated high-dimensional entangled states, we use quantum compressed sensing techniques [25]. Inspired by advanced classical methods for data analysis, these techniques significantly reduce the experimental cost for state reconstruction, are general for density matrices of arbitrary dimension, and have been experimentally demonstrated to characterise complex quantum systems [26, 27]. The general idea in compressed sensing, originally developed for classical data analysis, is to exploit the underlying structure that results from data collected in realistic situations to significantly reduce the number of parameters to be determined. In general, if the state to be characterised is low-rank (i.e. nearly pure), the approach requires quadratically fewer measurements than standard quantum tomography techniques, and the density matrix reconstruction from the data is efficient (requires solving a convex program). More details on the procedure are reported in appendix C.

We experimentally implemented compressed sensing QST to reconstruct the bipartite

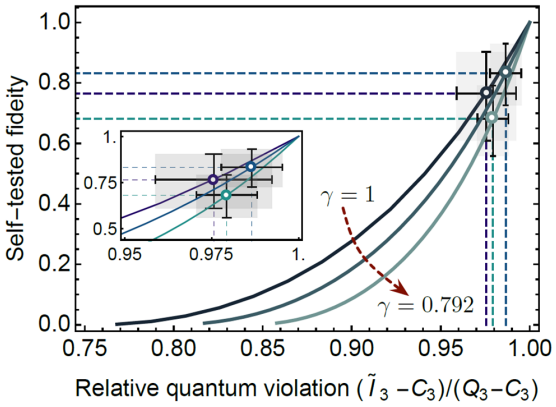


Figure 3.11: Self-testing of entangled qutrit states $(|00\rangle + \gamma\,|11\rangle + |22\rangle)/\sqrt{2+\gamma^2}$ for $\gamma = 1,\ 0.9,\ 0.792$. Self-tested minimal fidelities to ideal states are plotted as a function of the relative violation for more clarity. The significant uncertainty on the fidelity value is due to the general limited robustness of self-testing protocols.

entangled states generated with local dimension up to $d = 12$. Fidelities with ideal maximally entangled states $|\psi_d^+\rangle$ are reported in Fig. 3.10a. For dimensions $d = 4,\ 8$ and 12 the reconstructed density matrices are plotted in Fig. 3.10b-d, with fidelities of 96%, 87% and 81%, respectively.

3.2.4 *Self-testing of high-dimensional bipartite entangled states*

As a further characterisation test, and possible future application of the technology, we proceed by experimentally studying *self-testing* of entangled states in high dimensions. Widely studied in recent years, self-testing represents, broadly speaking, the characterisation of quantum devices without assumptions on their inner functioning, and is one of the practical applications on non-locality. The context of self-testing is the following. In a realistic implementation of quantum technologies, a company provides devices to some classical users which have no means to understand the complex inner functioning of the devices. If the users don't believe that the devices are reliable, and can thus make no assumptions about them, can they still certify if the devices are functioning in a correct way? That is, can they still characterise devices that they don't trust? Interestingly, non-locality allows us to do it. In particular, if the maximal violation of a Bell inequality can only be achieved by a unique quantum state and set of measurements, a near-optimal violation tells us that the state and measurements performed are close to the ideal ones, i.e. we characterise the entire experimental apparatus solely from the observed correlations. Recent theoretical developments allow us to make such statements quantitatively [28, 29]. In Ref. [23] it was shown that the SATWAP inequality can be used to self-test the maximally entangled state of two qutrits $|\psi_3^+\rangle$: employing a numerical approach from Ref. [29], based on the solution of a convex semi-definite program, a lower bound on the state fidelity can be obtained from the measured value of $\tilde{I}_3$. Such procedure can be also

generalised for arbitrary qutrit states of the form $|00\rangle + \gamma\,|11\rangle + |22\rangle\,/\sqrt{2+\gamma^2}$ (see appendix C). Fig. 3.11 shows how the self-tested fidelities lower bound change as a function of the observed violation of the SATWAP inequality relative to the classical bound, that is $(\tilde{I}_3 - C_3)/(Q_3 - C_3)$. Note that the the fidelity decreases very steeply with experimental noise, which means that very high precision is required to self-test non-trivial (> 0) lower bounds on the state fidelity. Despite this extreme noise sensitivity, experimental self-testing results for different values of $\gamma = 1,\ 0.9,$ and $(\sqrt{11} - \sqrt{3})/2 \approx 0.792$, reported in Fig.3.11, all presented good fidelities. In particular, we experimentally achieve an average self-tested fidelity of 77%.

Being possible only with such high level of precision, this application is a remarkable example of why it is important, when developing high-dimensional quantum technologies, to have the high-quality control enabled by our integrated quantum photonics approach.

3.3 Discussion

To summarise, the key points achieved in this experiment are:

- Ability to fabricate and control large-scale quantum photonic circuits, involving both the generation and processing of photonic states.

- High fidelities for the state generation and manipulation can be maintained even when scaling up the complexity of integrated photonic circuits.

- Demonstration of the potential of integrated quantum photonics for developing high-dimensional quantum technologies.

The on-chip generation, of a frequency-encoded high-dimensional entangled states was already demonstrated on a silicon nitride micro-ring frequency comb few months before we published our work (Ref. [4]). However, all the manipulation was performed off-chip, obtaining quite low fidelities. One of the novelties of this work is, therefore, to demonstrate that using integrated photonics actually brings important advantages compared to bulk high-dimensional approaches, showing the potential of silicon quantum photonics as a multidimensional platform. The level of precision and universality in controlling high-dimensional states demonstrated is unprecedented, and can open novel applications. For example, we showed for the first time the ability to self-test high-dimensional entangled states, a task which required high-fidelity in both the state generation and manipulation.

While the main scope of the experiment was to push the complexity of integrated optical circuits, the experiment was limited to the generation and processing of two

photons. It is, however, well-known that in a quantum computing architecture increasing only the dimensionality of a fixed number of photons in not a scalable approach: it is required that also the number of photons is increased. Generating states with a larger number of photons is, therefore, crucial for the development of large-scale integrated quantum photonics technologies. The next chapter is dedicated to this challenge.

Acknowledgements

The experiment was a strong collaboration between the author and Dr. Jianwei Wang, with additional technical help on the experiment from Dr. Raffaele Santagati, Dr. Davide Bacco, Dr. Damien Bonneau and Dr. Joshua W. Silverstone. The author's contributions are the mounting of the experimental set-up and the device, building up the hardware and software to control the silicon chip, and performing the experiment and the data analysis. The adaptation of the fiber optics scheme in Ref. [11] to generate high-dimensional entanglement in integrated photonics is due to Jianwei. The measurement scheme was designed by Dr. Anthony Laing. The silicon photonic device was fabricated by Dr. Yunhong Ding in the group of Prof. Karsten Rottwitt and Prof. Leif K. Oxenløwe at the Technical University of Denmark. For the quantum characterisation and applications, extremely helpful theoretical support was provided by Dr. Paul Skrzypczyk, Dr. Laura Mančinska from the University of Bristol, and Alexia Salavrakos, Dr. Jordi Tura, Dr. Remigiusz Augusiak, and Prof. Antonio Acín from the Institut de Ciencies Fotoniques (ICFO) in Barcelona. This work is published in Science (Ref. [1]).

References

1. Wang, J., Paesani, S., Ding, Y., Santagati, R., Skrzypczyk, P., Salavrakos, A., Tura, J., Augusiak, R., Mančinska, L., Bacco, D., *et al.* Multidimensional quantum entanglement with large-scale integrated optics. *Science* **350,** 285 (2018).
2. Krenn, M., Hochrainer, A., Lahiri, M. & Zeilinger, A. Entanglement by Path Identity. *Phys. Rev. Lett.* **118,** 080401 (2017).
3. Martin, A., Guerreiro, T., Tiranov, A., Designolle, S., Fröwis, F., Brunner, N., Huber, M. & Gisin, N. Quantifying photonic high-dimensional entanglement. *Phys. Rev. Lett.* **118,** 110501 (2017).
4. Kues, M., Reimer, C., Roztocki, P., Cortés, L. R., Sciara, S., Wetzel, B., Zhang, Y., Cino, A., Chu, S. T., Little, B. E., Moss, D. J., Caspani, L., Azaña, J. & Morandotti, R. On-chip generation of high-dimensional entangled quantum states and their coherent control. *Nature* **546,** 622–626 (2017).
5. Olislager, L., Safioui, J., Clemmen, S., Huy, K. P., Bogaerts, W., Baets, R., Emplit, P. & Massar, S. Silicon-on-insulator integrated source of polarization-entangled photons. *Opt. Lett.* **38,** 1960–1962. http://ol.osa.org/abstract.cfm?URI=ol-38-11-1960 (2013).
6. Silverstone, J. W., Bonneau, D., Ohira, K., Suzuki, N., Yoshida, H., Iizuka, N., Ezaki, M., Natarajan, C. M., Tanner, M. G., Hadfield, R. H., *et al.* On-chip quantum interference between silicon photon-pair sources. *Nature Photonics* **8,** 104 (2014).
7. Silverstone, J. W., Santagati, R., Bonneau, D., Strain, M. J., Sorel, M., O'Brien, J. L. & Thompson, M. G. Qubit entanglement between ring-resonator photon-pair sources on a silicon chip. *Nature Communications* **6,** 7948 (2015).

8. Paesani, S., Gentile, A. A., Santagati, R., Wang, J., Wiebe, N., Tew, D. P., O'Brien, J. L. & Thompson, M. G. Experimental Bayesian Quantum Phase Estimation on a Silicon Photonic Chip. *Phys. Rev. Lett.* **118,** 100503. https://link.aps.org/doi/10.1103/PhysRevLett.118.100503 (2017).

9. Wang, J., Paesani, S., Santagati, R., Knauer, S., Gentile, A. A., Wiebe, N., Petruzzella, M., O'Brien, J. L., Rarity, J. G., Laing, A. & Thompson, M. G. Experimental quantum Hamiltonian learning. *Nature Physics* **13,** 551–555 (2017).

10. Santagati, R., Wang, J., Gentile, A. A., Paesani, S., Wiebe, N., McClean, J. R., Morley-Short, S., Shadbolt, P. J., Bonneau, D., Silverstone, J. W., *et al.* Witnessing eigenstates for quantum simulation of Hamiltonian spectra. *Science Advances* **4,** eaap9646 (2018).

11. Schaeff, C., Polster, R., Lapkiewicz, R., Fickler, R., Ramelow, S. & Zeilinger, A. Scalable fiber integrated source for higher-dimensional path-entangled photonic quNits. *Opt. Express* **20,** 16145–16153 (2012).

12. Babazadeh, A., Erhard, M., Wang, F., Malik, M., Nouroozi, R., Krenn, M. & Zeilinger, A. High-Dimensional Single-Photon Quantum Gates: Concepts and Experiments. *Phys. Rev. Lett.* **119,** 180510. https://link.aps.org/doi/10.1103/PhysRevLett.119.180510 (2017).

13. Erhard, M., Fickler, R., Krenn, M. & Zeilinger, A. Twisted photons: new quantum perspectives in high dimensions. *Light: Science & Applications* **7,** 17146 (2018).

14. Krenn, M., Malik, M., Fickler, R., Lapkiewicz, R. & Zeilinger, A. Automated Search for new Quantum Experiments. *Phys. Rev. Lett.* **116,** 090405. https://link.aps.org/doi/10.1103/PhysRevLett.116.090405 (2016).

15. Carolan, J., Harrold, C., Sparrow, C., Martín-López, E., Russell, N. J., Silverstone, J. W., Shadbolt, P. J., Matsuda, N., Oguma, M., Itoh, M., Marshall, G. D., Thompson, M. G., Matthews, J. C. F., Hashimoto, T., O'Brien, J. L. & Laing, A. Universal linear optics. *Science* **349,** 711–716 (2015).

16. Brunner, N., Pironio, S., Acin, A., Gisin, N., Méthot, A. A. & Scarani, V. Testing the dimension of Hilbert spaces. *Phys. Rev. Lett.* **100,** 210503 (2008).

17. Sikora, J., Varvitsiotis, A. & Wei, Z. Minimum dimension of a Hilbert space needed to generate a quantum correlation. *Phys. Rev. Lett.* **117,** 060401 (2016).

18. Mermin, N. D. Simple unified form for the major no-hidden-variables theorems. *Phys. Rev. Lett.* **65,** 3373–3376 (1990).

19. Peres, A. Incompatible results of quantum measurements. *Phys. Lett. A* **151,** 107–108 (1990).

20. Brunner, N., Cavalcanti, D., Pironio, S., Scarani, V. & Wehner, S. Bell nonlocality. *Rev. Mod. Phys.* **86,** 419–478. https://link.aps.org/doi/10.1103/RevModPhys.86.419 (2014).

21. Wiseman, H. M., Jones, S. J. & Doherty, A. C. Steering, Entanglement, Nonlocality, and the Einstein-Podolsky-Rosen Paradox. *Phys. Rev. Lett.* **98,** 140402. https://link.aps.org/doi/10.1103/PhysRevLett.98.140402 (2007).

22. Cavalcanti, D & Skrzypczyk, P. Quantum steering: a review with focus on semidefinite programming. *Reports on Progress in Physics* **80,** 024001. http://stacks.iop.org/0034-4885/80/i=2/a=024001 (2017).

23. Salavrakos, A., Augusiak, R., Tura, J., Wittek, P., Acín, A. & Pironio, S. Bell inequalities tailored to maximally entangled states. *Phys. Rev. Lett.* **119,** 040402 (2017).

24. Collins, D., Gisin, N., Linden, N., Massar, S & Popescu, S. Bell inequalities for arbitrarily high-dimensional systems. *Phys. Rev. Lett.* **88,** 040404 (2002).

25. Gross, D., Liu, Y.-K., Flammia, S. T., Becker, S. & Eisert, J. Quantum state tomography via compressed sensing. *Phys. Rev. Lett.* **105,** 150401 (2010).

26. Gross, D. Recovering low-rank matrices from few coefficients in any basis. *IEEE Trans. on Information Theory* **57,** 1548–1566 (2011).

27. Bolduc, E., Gariepy, G. & Leach, J. Direct measurement of large-scale quantum states via expectation values of non-Hermitian matrices. *Nat. Commun.* **7,** 10439 (2016).

28. Mayers, D. & Yao, A. *39th Annual Symposium on Foundations of Computer Science (FOCS)* (1998).

29. Yang, T. H., Vértesi, T., Bancal, J.-D., Scarani, V. & Navascués, M. Robust and Versatile Black-Box Certification of Quantum Devices. *Phys. Rev. Lett.* **113,** 040401 (2014).

4 MULTI-PHOTON GENERATION AND INTERFERENCE ON A CHIP

In the previous chapter we have discussed how integrated optics, and in particular silicon photonics, represents a scalable approach to build large-scale photonic circuits, with hundreds of optical components linked together in optically stable interferometers. However, such experiments have so far been limited to the use of a small number of photons. For instance, in previous experiments no more than four photons coherently generated and processed on a chip have been reported [2]. In this experiment, using current technologies, we significantly improve this limit, generating and manipulating up to eight photons on a chip.

Increasing the number of photons is challenging due to the main sources of noise in linear optical quantum information: loss and non-deterministic photon generation. These errors are unique for quantum photonics, as in general the possibility that a qubit disappears is basically negligible in other platforms (solid-state qubits, trapped ions, etc.). For the ultimate goal of a photonic universal digital quantum computer, these noises will need to be corrected. However, the highly demanding overheads required for error tolerance will necessitate considerable technological progress before making such universal machines accessible. A nearer term goal is the development of non-universal devices that perform specialised algorithms; it is expected that a regime in which these devices can provide an advantage over conventional computing can be reached with significantly less error tolerance demands [3, 4]. In photonics, such a device would consist of a circuit where loss and non-deterministic photon sources are present, which can still generate and process a number of photons large enough to outperform classical approaches. The spirit of the experiment reported in this chapter is to push the integrated quantum photonics technology in this direction.

In particular, we will focus on a specific type of non-universal protocol for near-term

non-error-corrected photonic devices: boson sampling [3]. Apart from its importance as a viable route to demonstrate a quantum advantage, boson sampling can be mapped to useful applications, as we will see in the next chapter 5. Two different approaches to boson sampling are demonstrated and tested on the same chip, scattershot boson sampling and Gaussian boson sampling. The implementation allow us to benchmark the quality of the quantum interference for different number of emitted photons.

I will start by giving a brief background of boson sampling and its different flavours, and then go into the details of the experiment and results.

4.1 Background on boson sampling

The development of universal digital quantum computers represents one of the most important current technological challenges faced by the scientific community, and promise to unlock problems inaccessible with classical machines. However, the scalable implementation of complex algorithms on such universal quantum hardware is still far from accessible. In particular, the problem of fault-tolerant quantum computation is well beyond current technologies. On the other hand, there are already quantum systems currently accessible, e.g. ultra-cold atomic systems, that allow some degree of control and whose behaviour seems to be intractable to simulate on classical machines [5]. It is however difficult to interpret these systems in terms of computational machines, i.e. with some well-specified inputs and outputs.

Boson sampling has been proposed by Aaronson and Arkiphov as an intermediate situation: a well-defined computational model intractable on classical machines but realistic on near-term experimental capabilities [3]. The computational model is pictured in Fig. 4.1, and can be schematised as follows.

- **Inputs:** an initial multi-mode photonic state $|\psi\rangle$ and a linear-optical network with m input modes and m output modes, described by a $m \times m$ unitary U randomly sampled from the Haar distribution.
- **Experiment:** The input state $|\psi\rangle$ is propagated through the linear-optical circuit U and single-photon detection is performed on the evolved state, using m detectors on the output modes. (Here we will focus on single-photon detection, although the protocol can be further generalised to include Gaussian measurements as well [6]).
- **Outputs:** the observed n-photon coincidence pattern at the single-photon detectors, i.e. the combination of n detectors that produce a "click".

The computational model is a *sampling* problem: given the input state and the unitary, the experimenter is required to sample measurement outcomes from the resulting output

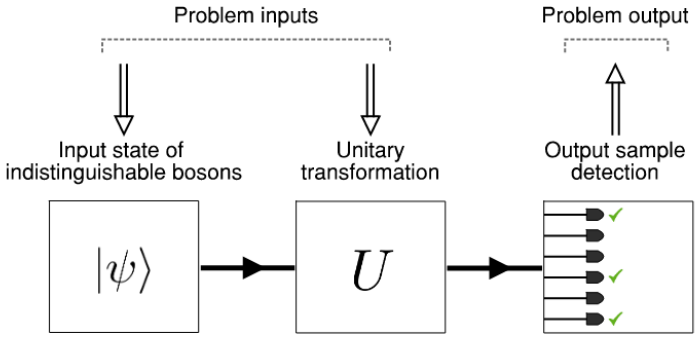


Figure 4.1: Schematic representation of the boson sampling computational problem.

distribution of the detection patterns. Clearly, this model is quite simple from a technological point of view; no feed-forward, optical non-linearity or adaptivity is required, making boson sampling machines much more realistic than fault-tolerant quantum machines for near-term experiments. In fact, boson sampling is arguably the simplest non-trivial type of quantum photonic experiment conceivable: prepare a state, evolve it through a linear-optical network, and measure the outcome. Still, quite surprisingly, it is intractable on classical computers if the number of photons is large enough ($n \gtrsim 50$ photons) [3, 7, 8]. Crucially, boson sampling is not only hard to simulate *exactly* on classical machines, but also *approximately* (under mild conjectures), meaning that the problem is classically intractable also if the quantum hardware is not fully error-corrected. The price to pay is universality: the boson sampling model is strongly believed to be non-universal, although interesting quantum simulation problems can be mapped into it, as is discussed in chapter 5 and Refs. [9–11].

The development of boson sampling has driven numerous experimental investigations in recent years from different quantum photonics groups [12–21]. As the first implementations were reported, it appeared clear that, despite the apparent experimental simplicity, with standard approaches to boson sampling it was a remarkable challenge for current photonic technologies to implement boson sampling in interesting regimes where some quantum speed-up could be achieved ($n \gtrsim 20$). So far demonstrations have been limited to $n \leq 5$ [20]. One of the main difficulties is the generation of large photonic states with non-deterministic sources. Following this difficulty, different approaches have been developed as more experimentally viable ways to increase the complexity of boson sampling experiments. The main difference between the various approaches is the choice of the input state $|\psi\rangle$, as is reviewed below.

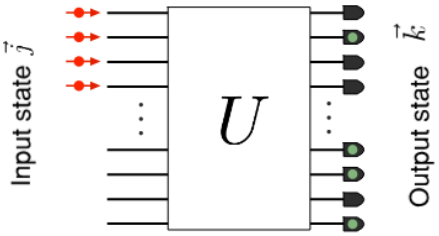


Figure 4.2: Schematic representation of the standard boson sampling protocol, where n photons are prepared and injected in a fixed set on n input ports of a $m \times m$ interferometer described by the unitary transformation U. Single photon detection is performed on the output m modes. The recorded detection patterns represent the outputs of the sampling algorithm.

4.1.1 *Standard boson sampling*

In its original proposal [3], boson sampling assumed as input state $|\psi\rangle$ a Fock state of n photons in n modes, with one photon per mode, as schematised in Fig. 4.2. Considering the case where the n photons are injected in the modes $\mathbf{j} = \{j_1, j_2, \ldots, j_n\}$ of a $m \times m$ interferometer (no more than a photon per mode), we can write the input state as

$$|\psi\rangle = |\mathbf{j}\rangle = \prod_{\ell=1}^{n} \hat{a}^\dagger_{j_\ell} |0\rangle \,, \tag{4.1.1}$$

where $\hat{a}^\dagger_j$ is the bosonic creation operator associated to the the optical mode j. As discussed in section 2.1, for non-interacting bosons, the action of the interferometer, described by the $m \times m$ unitary matrix U, on the input state is given by the transformation rule on the creation operators $\hat{\mathbf{a}}^\dagger \mapsto U^T \hat{\mathbf{a}}^\dagger$, so that the relation between the input creation operators ($\hat{a}^\dagger_j$) and output operators ($\hat{b}^\dagger_k$) is:

$$\hat{a}^\dagger_j = \sum_{k=1}^{m} U_{kj} \hat{b}^\dagger_k. \tag{4.1.2}$$

In boson sampling it is usual to consider output configurations $\mathbf{k} = \{k_1, k_2, \ldots, k_n\}$ where photons exit the interferometer in different modes, denoted as *collision-free* cases. This is because the "boson birthday paradox" states that in boson sampling scenarios, with $m = \mathcal{O}(n^2)$, collision-free cases happen with high probability [3, 22]. Under the transformation rule of eq.(4.1.2), in the collision-free scenario we have (see section 2.2.3):

$$\langle \mathbf{k}| \, U \, |\mathbf{j}\rangle = \mathrm{Perm}(U_{\mathbf{j},\mathbf{k}}), \tag{4.1.3}$$

where $U_{\mathbf{j},\mathbf{k}}$ is the submatrix of U obtained by taking its rows associated to $\mathbf{k}$ and columns associated to $\mathbf{j}$, as shown in Fig. 4.3a, and $\mathrm{Perm}(\cdot)$ is the permanent matrix function defined in eq.(2.2.17). Given the input Fock state $|j\rangle$, the probability to detect the output configuration $|k\rangle$ is then given by

$$p_U(\mathbf{k}|\mathbf{j}) = |\mathrm{Perm}(U_{\mathbf{j},\mathbf{k}})|^2. \tag{4.1.4}$$

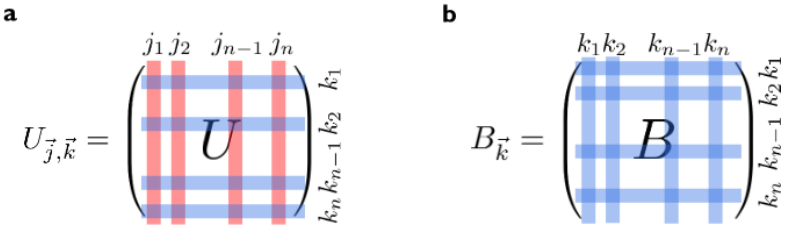


Figure 4.3: **a**, Calculation of the submatrix $U_{\mathbf{j},\mathbf{k}}$ for standard and scattershot boson sampling, where U is the unitary matrix associated to the interferometer. **b**, Calculation of the submatrix $B_{\mathbf{k}}$ for Gaussian boson sampling with vacuum squeezing, where $B = U[\oplus_i \tanh(\xi_i)]U^t$ and ξ_i are the input squeezing parameters.

Aaronson and Arkiphov have showed that, under mild conjectures, approximate sampling of output states $\mathbf{k}$ from the distribution $p_U(\mathbf{k}|\mathbf{j})$ is intractable on classical machines for large values of n. Current estimates predict that $n \approx 50$ are required to enter a regime where a classical simulation of boson sampling would no longer be possible on supercomputers [7].

While the impossibility to classically approximate a boson sampler indicates some sort of robustness of the protocol against experimental imperfections, it is not clear how noise affects the complexity of boson sampling [23]. A particularly challenging type of imperfection is the lack of ideal deterministic single photon sources. In fact, if n sources with an efficiency ϵ are used to produce the n-photon input Fock state in the standard boson sampling configuration, the probability of generating such state is proportional to ϵ^n. Therefore, if the sources are non-deterministic ($0 \leq \epsilon < 1$), the experimenter would need to wait a time that increases exponentially with n before observing an event. Although progress has been reported with solid-state single photon emitters, current single photon sources are limited to system efficiencies of $\epsilon \lesssim 0.3$ (see for example Ref. [20]), which, considering no losses in the rest of the circuit and perfect detection, would practically restrict implementations to approximately 20 photons. This limitation has motivated the development of more experimentally viable alternatives [24–29]. Two prominent approaches are scattershot [30] and Gaussian boson sampling [31]. A key feature of these two protocols is to remain efficient even in presence of imperfect sources.

4.1.2 *Scattershot boson sampling*

The scattershot approach to boson sampling, developed by Lund et al. few years after the original boson sampling proposal [30], represents a way to increase the complexity of photonic experiments with realistic non-ideal sources, such as those based on spontaneous processes (i.e. SPDC or SFWM sources). The protocol is represented in Fig. 4.4.

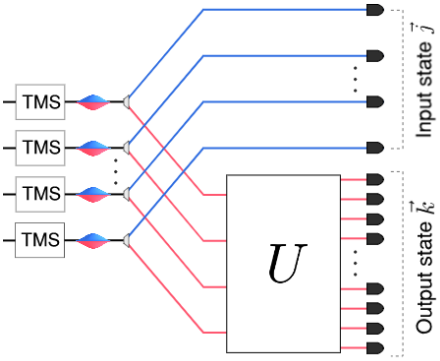


Figure 4.4: Schematic of scattershot boson sampling. Weak two-mode squeezed states are generated and separated. The idlers modes are measured to herald the modes in which the signal photons enter the interferometer, i.e. the detection pattern $\mathbf{j}$ measured in the idler modes corresponds to the input state of the signal photons, while the output pattern $\mathbf{k}$ is detected after the signals have interfered. For this protocol, which consists in sampling from the joint distribution $p_U(\mathbf{k}, \mathbf{j})$, the only input is the unitary U, while the outputs samples are the joint detection patterns $(\mathbf{k}, \mathbf{j})$

To perform n-photon boson sampling, $m_0 \geq n$ parametric sources, each generating a two-mode squeezed vacuum state (TMS)

$$|\mathrm{TMS}(\xi)\rangle_j = \frac{1}{\cosh(\xi)} \sum_{n=0}^{\infty} \tanh(\xi)^n \, |n\rangle_{i_j} \, |n\rangle_{s_j} \,, \tag{4.1.5}$$

where ξ represents the two-mode squeezing parameter and i_j (s_j) the mode associated to the idler (signal) photons for the j-th source. The total state generated is then simply a product of TMS states produced from the array of sources

$$|\psi\rangle = \bigotimes_{j=1}^{m_0} |\mathrm{TMS}(\xi)\rangle_j \,, \tag{4.1.6}$$

where for simplicity we assume the same squeezing ξ in all sources. We can already note a difference with standard boson sampling: the input resource state. Here we use two-mode squeezed states, deterministically generated in spontaneous processes, rather than a fixed Fock state with single photons. After the state generation, all the idler modes are directly sent to photon counters, while the $m_0 \leq m$ signal modes are injected into an $m \times m$ interferometer described by U. Suppose now that in the idler modes a collision-free configuration with n single photons in the modes $\mathbf{j} = \{j_1, j_2, \ldots, j_n\}$ (and no photons in the remaining $m_0 - n$ modes) is detected. Then, due to the perfect correlations of the photon number in eq.(4.1.5), also the state of the signal photons, before entering the interferometer, collapses in the same configuration $|\mathbf{j}\rangle_s$. We thus recover a standard boson sampling scenario where the input state $|\mathbf{j}\rangle$ is heralded by the detected pattern in the idler modes. In fact, the probability of generating a particular n-photon input state $|\mathbf{j}\rangle$ is

$$p(|\mathbf{j}\rangle) = \tanh(\xi)^{2n} \, \mathrm{sech}(\xi)^{2m_0}, \tag{4.1.7}$$

which, as for standard boson sampling, decreases exponentially with n.

However, the idea of scattershot boson sampling is to not limit the experiment to a particular input configuration, but to accept the input whenever n of the m_0 idler detectors click. This corresponds to a boson sampling experiment where the input state for each runs is chosen randomly (but is known from the heralding) between all the possible configurations $|\mathbf{j}\rangle$. Scattershot boson sampling thus consists of a random (but heralded) Fock state generator obtained by collapsing the state of multiple two-mode squeezers via detecting the idler photons, and then injecting the signal photons into the unitary. Crucially, the probability of generating a random state from probabilistic sources is significantly higher than the probability of a specific input. In fact, as the number of configurations with n photons between the m_0 idler modes (one photon per mode) is $\binom{m_0}{n}$, the probability to generate an n-photon input state is now combinatorially enhanced to

$$p(n,\xi,m_0) = \binom{m_0}{n} \tanh(\xi)^{2n} \operatorname{sech}(\xi)^{2m_0}. \tag{4.1.8}$$

The input/output distribution is then simply given by $p_U(\mathbf{k},\mathbf{j}) = p_U(\mathbf{k}|\mathbf{j})p(\mathbf{j})$, where the transition probability $p_U(\mathbf{k}|\mathbf{j})$ is the same as in standard boson sampling (eq.(4.1.4)) and $p(\mathbf{j})$ is the uniform random distribution[1] over all possible input configurations $\mathbf{j}$. Demonstrating that such a randomised form of boson sampling is as hard as the standard approach for classical machines is relatively straightforward[2] [30].

In the typical regime for boson sampling we consider $m_0 = m = n^2$, for which the generation probability is

$$p(n,\xi) = \binom{n^2}{n} \tanh(\xi)^{2n} \operatorname{sech}(\xi)^{2n^2}. \tag{4.1.9}$$

In this regime, as shown in Fig. 4.5, the optimal probability to generate an n-photon state, calculated by fixing n in eq.(4.1.9) and maximising over ξ [30], is given by:

$$p_{\max}(n) \simeq \frac{1}{e\sqrt{2\pi n}}, \tag{4.1.10}$$

1 A uniform distribution between all the input configurations is obtained assuming equal squeezing in all sources. If the squeezing varies, the distributions would no longer be uniform.

2 Intuitively, the proof reduces to the fact that each instance of scattershot is equivalent to a run in the standard boson sampling scenario but with a randomised input. That means that the columns used to obtain the submatrix $U_{\mathbf{j},\mathbf{k}}$ in eq.(4.1.4) are no longer fixed and can change at every run. However, as the matrix U is Haar random, the statistical properties of the submatrices $U_{\mathbf{j},\mathbf{k}}$ would again satisfy the same conditions as for the standard approach. Therefore, the arguments by Aaronson and Arkiphov for the original proof of hardness still hold for the scattershot approach.

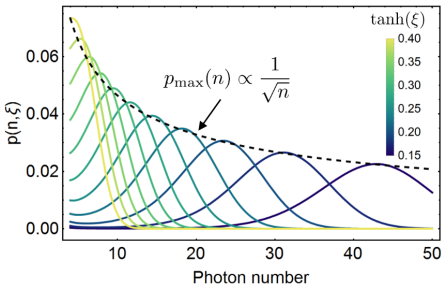


Figure 4.5: Probability of generating n photon input state from n^2 sources of two-mode squeezing ξ in the scattershot regime. Darker colours correspond to lower squeezing values, as shown in the legend. The black dashed line represents the optimal probability as a function of n (eq.(4.1.10)), which presents a polynomial scaling $1/\sqrt{n}$.

which is obtained for

$$\tanh(\xi_{max}) = \frac{1}{\sqrt{n+1}}. \tag{4.1.11}$$

An important result can be drawn from eq.(4.1.10): the generation probability in scattershot decreases only polynomially ($\propto n^{-1/2}$) when increasing the number of photons, in contrast to an exponential decrease in the standard regime. This means that scattershot boson sampling can be efficient also in presence of non-deterministic sources with a finite level of squeezing. Actually, from eq.(4.1.11), the optimal squeezing gets lower for higher photon numbers n.

Compared to standard boson sampling, such an important improvement comes at the cost of an increase in the required number of sources (n^2 against n), and a higher number of detectors (twice as many). However, we have already seen in chapter 3 that the implementation of large arrays of sources based on spontaneous processes is possible in integrated quantum circuits. The integration of superconducting detectors [32] or time-multiplexing schemes [33] can be viable approaches to gain access to hundreds of efficient detectors. There are thus prospects for large-scale implementations of scattershot-type experiments in integrated optics.

In general, the scattershot approach could be seen as a type of error correction: in the presence of experimental imperfections, it maintains the scaling of a protocol efficient by introducing overheads in the hardware. Note that the error addressed here is the imperfection in the sources, but other errors due, for example, to losses in the circuit still represent important issues also for scattershot. With no correction for losses, also scattershot boson sampling at some point becomes inevitably non-scalable. However, it greatly pushes forward such limit compared to the standard approach. The hope is that, supported by technological improvement in terms of losses, the scattershot approach could allow us to push the threshold beyond the regime ($n \gtrsim 20$) where interesting applications of boson sampling are possible.

The idea behind scattershot boson sampling was to change the initial resource state used. In particular, we passed from Fock states, that are highly non-Gaussian, to

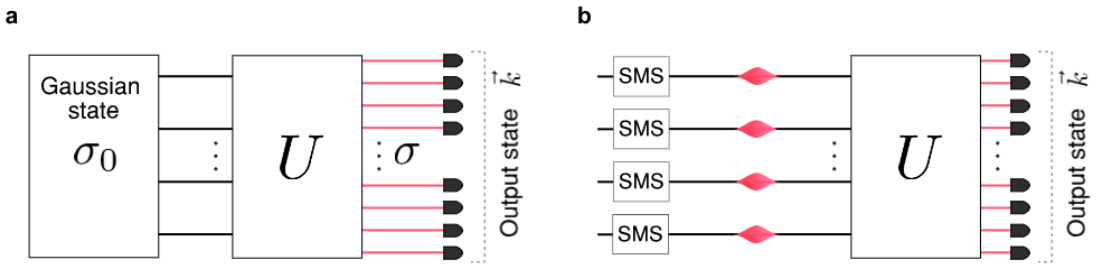


Figure 4.6: **a**, Schematic of Gaussian boson sampling. A multi-mode Gaussian state described by a covariant matrix σ_0 in generated and injected into the interferometer U. The Gaussian state emerging from the unitary has covariance σ, from which a photon-number configuration with n photons is obtained via single photon detection at the output. **b**, Gaussian boson sampling using multiple sources of single-mode squeezing at the input.

Gaussian TMS states. An advantage of some Gaussian states, e.g. squeezed states, is that they can be produced deterministically in spontaneous processes, while n-photon state generation is always probabilistic. While in scattershot this aspect is exploited to increase the generation rate, the state is collapsed to a photon-number state before the interference via the heralding measurement, so that the state injected in the interferometer is actually a photon-number one, and the interference is described by the same statistics as in standard boson sampling (see eq.4.1.4). An interesting question that has been investigated in recent years is how the description (and the computational complexity) would change in the case where the collapse into a photon number state happens only after the interference, i.e. with Gaussian states directly fed into the interferometer.

4.1.3 *Gaussian boson sampling*

Boson sampling with Gaussian states is an interesting generalisation of the scattershot protocol, which didn't fully explore the Gaussian nature of the input resource as it was collapsed to a photon-number state before the interference. In the protocol of Gaussian boson sampling, input Gaussian states are directly injected into the interferometer and the state is collapsed via single photon counters only after the interference [31, 34]. From an experimental perspective, compared to scattershot, this could be a more resource-efficient approach as all generated photons are interfered, while no photons and no additional detectors are needed for heralding. Moreover, novel applications have been proposed that are based on boson sampling with Gaussian states [9, 35].

The general protocol of Gaussian boson sampling is shown in Fig. 4.6a. The ingredients needed are an input Gaussian state, a unitary transformation U, and single photon detection at the output. As discussed in section 2.1, Gaussian states allows a simple

analytical description in terms of the first two moments of their Gaussian distributions in the parameter space, i.e. the displacement $\mathbf{d}$ and the covariance matrix $\boldsymbol{\sigma}$. For simplicity we here consider only states with zero displacement (e.g. vacuum-squeezed states), which allow us to describe the initial m-mode Gaussian state generated in the boson sampling protocol via its $2m \times 2m$ covariance matrix $\boldsymbol{\sigma}_0$. When a Gaussian state evolves through a linear interferometer U, the covariance matrix of the output state $\boldsymbol{\sigma}$ can be obtained via the symplectic transformation (see eq. 2.1.33)

$$\boldsymbol{\sigma} = \begin{bmatrix} U & 0 \\ 0 & U^* \end{bmatrix} \boldsymbol{\sigma}_0 \begin{bmatrix} U^\top & 0 \\ 0 & U^\dagger \end{bmatrix}. \tag{4.1.12}$$

The last ingredient we now need to describe the scheme is the photon-counting from the m-mode state $\boldsymbol{\sigma}$, that is, the probability to collapse to a n-photon state $|\mathbf{k}\rangle$ when measuring the output state with single-photon detectors. In Ref. [31] the authors formulated a closed form for such probability, given by

$$p_{\boldsymbol{\sigma}}(\mathbf{k}) = \frac{1}{n_{\mathbf{k}}!\sqrt{\det(\boldsymbol{\sigma}_Q)}} \mathrm{Haf}(A_{\mathbf{k}}), \tag{4.1.13}$$

where $n_{\mathbf{k}}! = n_1(\mathbf{k})!n_2(\mathbf{k})!\ldots n_m(\mathbf{k})!$, and $n_i(\mathbf{k})$ is the number of photons in the j-th mode for the configuration $\mathbf{k}$. For collision-free configurations we thus have $n_{\mathbf{k}}! = 1$. The matrices $\boldsymbol{\sigma}_Q$ and A are given by

$$\boldsymbol{\sigma}_Q = \boldsymbol{\sigma} + \mathbb{1}_{2m}, \qquad A = \begin{bmatrix} 0 & \mathbb{1}_m \\ \mathbb{1}_m & 0 \end{bmatrix} \left(\mathbb{1}_{2m} - \boldsymbol{\sigma}_Q^{-1}\right), \tag{4.1.14}$$

and $A_{\mathbf{k}}$ is the submatrix of A obtained by selecting its $\{k_1, m+k_1, k_2, m+k_2, \ldots, k_n, m+k_n\}$ rows and columns [31]. The Hafnian function of a $2m \times 2m$ matrix

$$\mathrm{Haf}(X) = \sum_{\mu \in C_{2m}} \prod_{i=1}^{m} X_{\mu(2i-1),\mu(2i)}, \tag{4.1.15}$$

with C_{2m} the set of canonical permutations on $2m$ elements, is a generalisation of the permanent function, firstly introduced by Caianiello in the 70s [36]. In particular, the permanent of a matrix can be expressed in terms of the Hafnian via

$$\mathrm{Perm}(X) = \mathrm{Haf}\left(\begin{bmatrix} 0 & X \\ X^\top & 0 \end{bmatrix}\right). \tag{4.1.16}$$

Moreover, for the Hafnian function it holds that [36]

$$\mathrm{Haf}(X_1 \oplus X_2) = \mathrm{Haf}(X_1)\mathrm{Haf}(X_2). \tag{4.1.17}$$

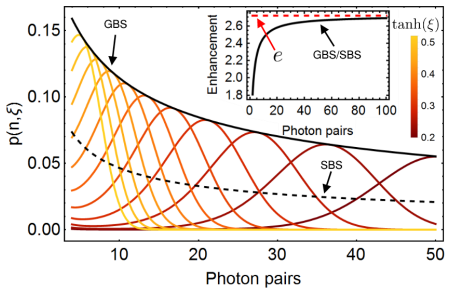


Figure 4.7: Probability of generating input state with n photon pairs ($2n$ total photons) from n^2 sources with single-mode squeezing ξ in the Gaussian boson sampling regime. Darker colours correspond to lower squeezing values, as shown in the legend. The black solid line represents the optimal probability as a function of n. As shown the inset, the relative speed-up respect to the scattershot case (black dashed line in the main plot) tends to a constant value of $e(\simeq 2.71)$ asymptotically.

While eq.(4.1.13) is very elegant and valid for arbitrary input Gaussian states, it may still be quite intricate to calculate. The situation is greatly simplified if we consider single-mode squeezed states at the input, as shown in Fig. 4.6b, on which we focus in the remaining of the chapter. In this scenario, each source generates the state

$$|\mathrm{SMS}(\xi)\rangle_j = \frac{1}{\sqrt{\cosh \xi}} \sum_{n=0}^{\infty} (-\tanh \xi_k)^n \frac{\sqrt{(2n)!}}{2^n n!} |2n\rangle_{s_j}\,, \tag{4.1.18}$$

and the total state is given by

$$|\psi\rangle = \bigotimes_{j=1}^{m_0} |\mathrm{SMS}(\xi)\rangle_j\,. \tag{4.1.19}$$

The covariance matrix $\boldsymbol{\sigma}_0 = \frac{1}{2} S(\boldsymbol{\xi}) S^\dagger(\boldsymbol{\xi})$ is obtained from the squeezing symplectic matrix

$$S(\boldsymbol{\xi}) = \begin{bmatrix} \oplus_{i=1}^m \cosh(\xi_i) & \oplus_{i=1}^m \sinh(\xi_i) \\ \oplus_{i=1}^m \sinh(\xi_i) & \oplus_{i=1}^m \cosh(\xi_i) \end{bmatrix}, \tag{4.1.20}$$

with ξ_i the single-mode squeezing parameter for the source in mode i. The output state after the interferometer in Fig. 4.6b has, therefore, covariance

$$\boldsymbol{\sigma} = \frac{1}{2} \begin{bmatrix} U & 0 \\ 0 & U^* \end{bmatrix} S(\boldsymbol{\xi}) S^\dagger(\boldsymbol{\xi}) \begin{bmatrix} U^\top & 0 \\ 0 & U^\dagger \end{bmatrix}, \tag{4.1.21}$$

for which the matrix A in eq.(4.1.14) can be simply written as a direct sum $A = B(\boldsymbol{\xi}) \oplus B^*(\boldsymbol{\xi})$, with [31, 34]

$$B(\boldsymbol{\xi}) = U \left(\bigoplus_{i=1}^m \tanh(\xi_i) \right) U^\top. \tag{4.1.22}$$

From eq.(4.1.13) and eq.(4.1.17), in the collision-free we can thus finally write:

$$p_{\boldsymbol{\xi}}(\mathbf{k}) = \frac{1}{\sqrt{\det(\boldsymbol{\sigma}_Q)}} \left|\mathrm{Haf}\left[B_{\mathbf{k}}(\boldsymbol{\xi})\right]\right|^2, \tag{4.1.23}$$

where the submatrix $B_{\mathbf{k}}$ is obtained from the rows and columns $\{k_1, k_2, \ldots, k_n\}$ of B, as pictured in Fig. 4.3b.

The task of Gaussian boson sampling with single-mode squeezed states as inputs can thus be summarised as the task of sampling output configuration $\mathbf{k}$ from the distribution $p_{\boldsymbol{\xi}}(\mathbf{k})$ in eq.(4.1.23). Moreover, as the Hafnian encapsulates the permanent function via eq.(4.1.16), in the case of single-mode vacuum squeezing evidence was provided that also Gaussian boson sampling is still computationally hard on classical computers, both in the exact and approximate case [31, 34]. However, as the computation time for the calculation of a permanent and a Hafnian of a $n \times n$ matrix are $\mathcal{O}(n2^n)$ and $\mathcal{O}(2^{n/2})$ respectively [37], it is expected that $2n$ photons are required for a Gaussian boson sampling protocol to achieve a classical run time comparable with a n photon standard boson sampling experiment.

From an experimental point of view, Gaussian boson sampling is an interesting resource efficient alternative compared to previous approaches. First, it provides an enhancement in terms of generation rate. In particular, the probability of generating n pairs of photons ($2n$ photons in total) from m_0 single-mode squeezing sources is given by

$$p(n, \xi, m_0) = \binom{n + m_0/2 - 1}{n} \tanh(\xi)^{2n} \,\mathrm{sech}(\xi)^{m_0}. \tag{4.1.24}$$

In Fig. 4.7 the generation probability is shown for different numbers of photons, using $m_0 = n^2$. As for the scattershot approach, the optimal generation probability decreases as $\propto 1/\sqrt{n}$. A comparison with the generation rate obtained from scattershot boson sampling can also be drawn. An improvement for the Gaussian approach can be observed from the inset of Fig. 4.7, with the relative rate asymptotically converging to a factor $e \simeq 2.71$.

Moreover, as no heralding is performed, the number of detectors required is halved compared to scattershot. Additionally, in Ref. [31, 34] the authors also suggest that the computational complexity of Gaussian boson sampling for an n-photon experiment should be saturated by using approximately n sources, which, compared to scattershot, would represent a quadratic resource advantage. Note also that, as no photons are used for heralding, twice as many photons would participate in the interference.

4.1.4 *Validation of boson sampling*

An important challenge in boson sampling is the task of verifying that the experiment is sampling from a probability distribution arising from multi-photon quantum interference, rather than from some other distribution which could be efficiently simulatable by a classical computer. While it is thought that in the general case there is no scalable approach to verify a boson sampler, different methods have been proposed to gather supporting evidence for the correct implementation of the protocol [38].

The task of boson sampling validation can be interpreted as the problem of estimating a confidence on the sampler behaving accordingly to an ideal model $\mathcal{M}_0$. A powerful tool to tackle such task is the use of Bayesian model comparison techniques [39]. These can be generalised to validate both scattershot and Gaussian boson sampling, and we will thus use it as a consistent validation approach for the different protocols that will be reported in the next section.

The procedure is as follows. Suppose that, given a set of data $D = \{x_i\}_{i=0}^{N}$ consisting of N output samples x_i from the interferometer, we want to verify if the data is more likely to arise from an ideal model $\mathcal{M}_0$ or from a test model $\mathcal{M}_t$. For example, $\mathcal{M}_0$ can represent an ideal implementation of scattershot or Gaussian boson sampling, while $\mathcal{M}_t$ a classical implementation where all photons are distinguishable. Given the set of samples D collected in the experiment, Bayes' rule offers an immediate way to estimate the confidence in model $\mathcal{M}_0$, that is the probability for model $\bar{\mathcal{M}}$ to represent the underlying experiment, which is given by:

$$\begin{aligned} p(\mathcal{M}_0|D) &= \frac{p(D|\mathcal{M}_0)p_0(\mathcal{M}_0)}{p(D|\mathcal{M}_0)p_0(\mathcal{M}) + p(D|\mathcal{M}_t)p_0(\mathcal{M}_t)} \\ &= \frac{1}{1 + \frac{p(D|\mathcal{M}_t)p_0(\mathcal{M}_t)}{p(D|\mathcal{M}_0)p_0(\mathcal{M}_0)}}. \end{aligned}$$

Assuming statistical independence between different events, the probability $p(D|\mathcal{M}_0)$ is given by $p(D|\mathcal{M}_0) = \prod_{i=1}^{N} p(x_i|\mathcal{M}_0)$, where $p(x_i|\mathcal{M}_0)$ is the probability of obtaining the measured outcome x_i according to the ideal model (dictated by the permanent function for scattershot and by the Hafnian in Gaussian boson sampling), and similarly for $p(D|\mathcal{M}_t)$.

Note that, in the case where the detection is restricted to collision-free events, the probabilities $p(x_i|\mathcal{M}_0)$ have to be normalised to take it into account, which corresponds to dividing by the sum of the probabilities over all collision-free events. The reason for this is that the probability of a sample being collision-free differs between the different models. Therefore different normalisation constants have to be used for the two models tested.

The probabilities $p_0(\mathcal{M}_0) = p_0(\mathcal{M}_t) = 0.5$ represent a prior distribution on the true model, which we here assume to be uniform to avoid any bias. The confidence $p(\mathcal{M}_t|D)$ for the adversary model $\mathcal{M}_t$ can be calculated in the same way. Note that, as the approach requires to estimate probabilities $p(x_i|\mathcal{M}_0)$, which is not efficient for classical machines, the approach is not scalable. However, it is general and can be used to validate against any model which allows us to calculate probabilities $p(x_i|\mathcal{M}_t)$.

4.2 Experimental set-up and photonic chip

While large realisations of boson sampling protocols will require a complexity which is likely to be achievable only via the integration of photonic componentry, the few demonstrations reported to date have relied heavily on bulk optical elements [13–16, 18–20, 33, 40]. One of the scopes of the experiment reported here is take a technological step forward, and achieve boson sampling experiments where states with many photons are generated and interfered on a silicon photonic chip.

4.2.1 *Photonic circuit and experimental set-up*

The silicon photon device used is shown in Fig. 4.8a. While most of the components and the fabrication procedures are essentially identical as for the device reported in chapter 3, an important difference is the integration of ultra-low loss chip-to-fibre couplers [41]. These comprised an apodised grating coupler assisted by an aluminium mirror underneath it [42, 43]. This technology allowed us to collect photons with $\approx 80\%$ efficiency from the chip, which played a great role in increasing the photon number.

Input quantum states are generated in an array of four integrated spiral sources via SFWM. The sources are coherently pumped by splitting the input laser via MMIs. To implement different boson sampling protocols, the sources are operated in two different regimes. In the first regime we pump them using a single-wavelength laser in order to generate weak two-mode squeezed vacuum (TMS) states via non-degenerate SFWM. In this case, signal and idler photons are emitted at different wavelengths[3]. In the second regime, we use a dual-wavelength pumping scheme and generate weak single-mode squeezed vacuum (SMS) states via degenerate SFWM. In this second case all photons are emitted at the signal frequency. Both regimes have been discussed in section 2.4. The switching between the two different regimes is performed off-chip, as described in the following section. After the sources, the pump light is filtered out of the circuit

3 The signal (idler) wavelength used was 1549.3 (1536.6) nm. To perform non-degenerate SFWM a single pump at 1542.9 nm (1.6 nm bandwidth) was used, while in the dual-pump scheme for degenerate SFWM two spectral slices centred at 1552.5 nm and 1549.3 nm were used (again with 1.6 nm bandwidth).

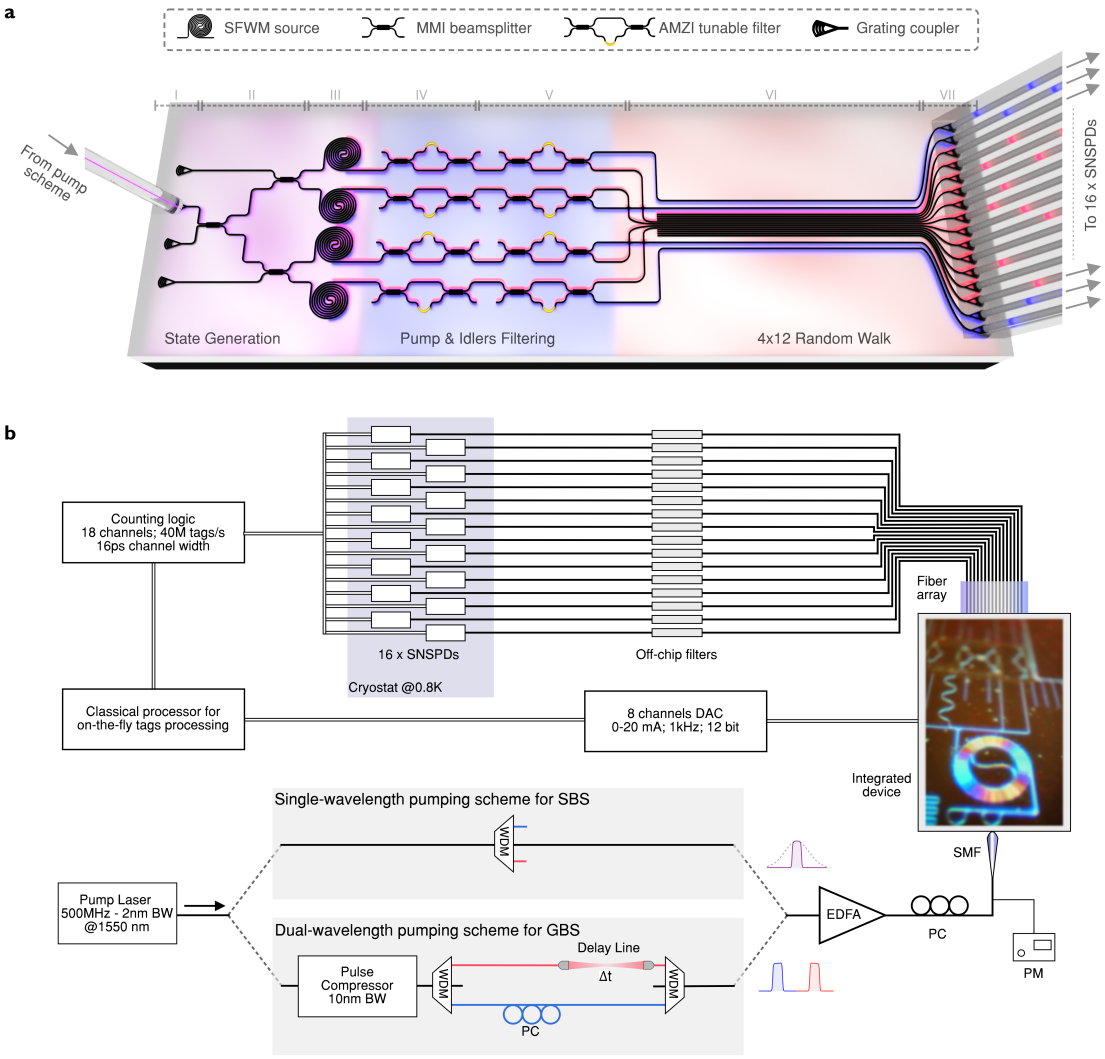


Figure 4.8: **a**, Chip schematic. Input pump laser light (purple) is injected into the chip (I), split via a network of MMI beam-splitters (II), and used to coherently pump an array of four SFWM sources (III). After the sources, the pump is filtered out using AMZI on-chip filters (IV). The idler (blue) photons are separated from the signal (red) photons via AMZIs (V). Signal photons undergo a quantum random walk obtained coupling 12 waveguides (VI). Finally, all photons are collected via a fibre array (VII) and sent for detection to an array of 16 SNSPDs. **b**, Schematic of the full experimental apparatus, including the apparatus to switch between the two pumping schemes.

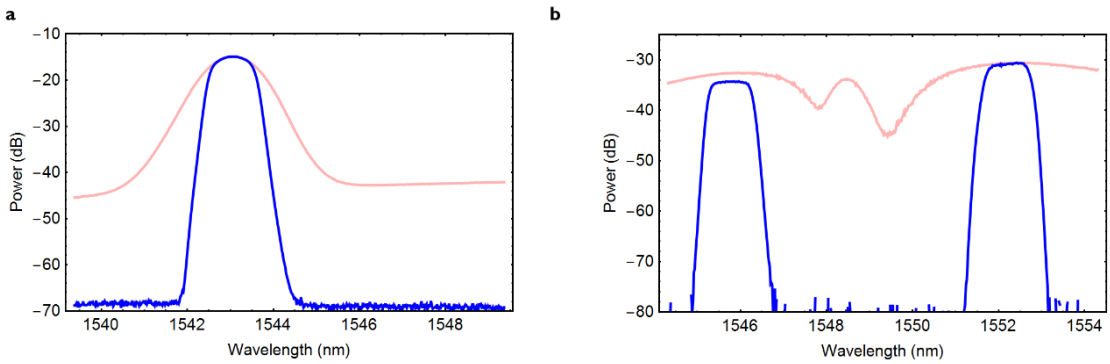


Figure 4.9: **a**, Spectrum of the input pump light in the single-wavelength pumping scheme, measured via an optical spectrum analyser. The red opaque line is the measured spectrum emitted by the laser, while the blue line is the spectrum of the light injected into the chip after being filtered. **b**, Spectrum of the input pump light in the dual-wavelength regime. The red opaque line is the spectrum emerging from the pulse compressor. The blue line is the spectrum of the light injected into the chip after selecting and combining two spectral slices via the use of WDMs.

via integrated AMZI filters to avoid generation of spurious photons in the rest of the chip. A second layer of AMZIs was used to separate the idler photons, used to perform the heralding in scattershot boson sampling. The signal photons are routed into the four central modes of an integrated continuous quantum random walk, obtained by evanescently coupling together 12 waveguides. Such structure allows us, due to phase fluctuations, to implement a 4×12 random low-loss interferometer for boson sampling experiments [44, 45]. After the interference, all photons, including the idlers, are collected out of the chip via a fibre array through the low-loss couplers, routed to superconducting nano-wire single photon detectors (SNSPDs), as shown in Fig. 4.8b. In total an array of 16 SNSPDs was used to detect the 4 heralding modes and the 12 outputs of the interferometer.

To support the high count rates observed in the experiment, we installed a fast counting logic able to support up to $40 \cdot 10^6$ events/second between the 16 detectors. The counting logic was interfaced with a customised computer able to process the data on the fly [4].

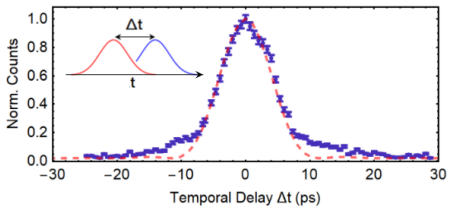


Figure 4.10: Coincidence counts emitted in the degenerate SFWM regime for different delays between the two pumps with different wavelengths. Error bars are obtained from Poissonian photon statistics.

4.2.2 *Tunable pumping schemes*

To switch between degenerate and non-degenerate SFWM regimes, required to operate Gaussian boson sampling and scattershot boson sampling, respectively, we devised an off-chip scheme to spectrally shape the injected pump accordingly. The scheme is reported in Fig. 4.8b.

To perform non-degenerate SFWM, producing two-mode squeezing, a simple single-pump scheme was required. This was easily achieved using an external laser emitting transform-limited pulses with a 2 nm bandwidth and 500 MHz repetition rate. The laser was filtered off-chip to remove spurious photons in the spectral tails of the laser, optically amplified, and finally injected into the chip. To generate single-mode squeezing via degenerate SWFM a more complex dual-pump scheme is instead needed, where two pulses with different wavelengths have to arrive simultaneously at the sources. A possible way to obtain a pump light with such properties could be to combine two different lasers, one for each frequency. While such approach would be straightforward to implement with continuous-wave lasers [46], with pulsed pump light it would require complex clock synchronisation, phase locking, and good spectral matching between the two lasers. We therefore opted for a different approach, represented in Fig. 4.8b. We used a pulse compressor to spectrally broaden (i.e. temporally compress) the pulses emitted from the laser to a total bandwidth of approximately 10 nm, from which two spectral slices are selected via a wavelength-division de-multiplexer (WDM), as shown in 4.9b. The two selected spectral regions are then recombined into a single fibre via a second WDM. A delay line between the two WDMs is used to ensure the temporal overlap of the two pulses with different frequencies in the sources. As can be observed from Fig. 4.10, a peak of detected photon events in the signal wavelength is observed when the two different pulses are temporally overlapped, due to the photon generation in the degenerate SFWM regime. A low level of noise can be observed when pulses are not overlapped[5].

4 Without the on-line data processing, we would have needed to save approximately 1 GByte/minute of time tags, which would have made the data collection impractical for long measurements.

5 The noisy background events detected when the two pulses are not temporally overlapped are mainly due to multi-pair emissions via non-degenerate SFWM from the two different pumps. A signal-to-noise ratio of ≈ 30 can be estimated from Fig. 4.10.

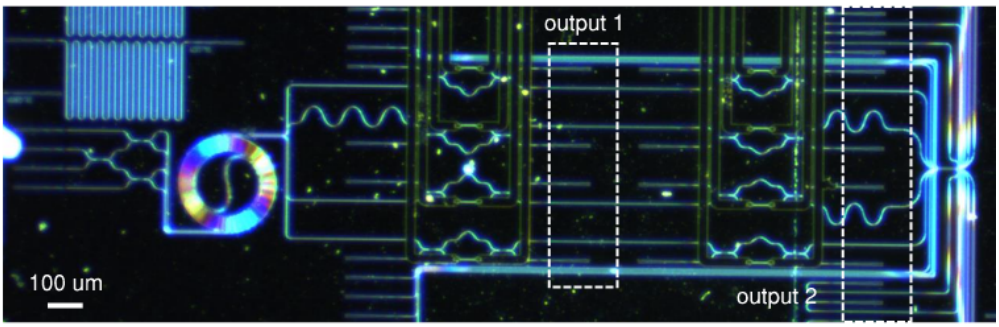


Figure 4.11: Optical microscope image of the integrated quantum photonic circuit. On the left the four sources, wrapped together in a spiral shape, can be observed, where photons are generated after the pump is coherently split by the MMI beam-splitters on the far left side. In the central region are the two layers of AMZI filters, with the electrical wiring used to control the thermal phase-shifters. On the right the continuous random walk, obtained by evanescently coupling 12 waveguides for a length of 110 μm, can be observed. The 12 outputs of the random walk and the 4 idler modes after the filters are routed to an array of 16 grating couplers (output 2), where photons are collected out of the chip via a fibre array (which was removed while taking this picture in order to show the photonic circuit below it). Output 1, which allows us to collect photons generated from each single source before the interferometer, can instead be used for characterising the sources.

The pump spectra in the two different pumping schemes are reported in Fig. 4.9. It can be noted that in the dual-pump scheme a large part of the pump power is lost when selecting the two slices from the broadband spectrum. Practically, this constrained us to use significantly less pump power when operating in the degenerate SFWM regime.

4.2.3 *Characterisation of the experiment*

The fabrication of a low-loss device is a crucial technological step required for multi-photon experiments with integrated photonics. The channel efficiency, that is the total loss experienced by each photon from generation to detection, can be measured via the photon coincidences to singles ratio $\eta_i = C_i/CC$, where C_i is the singles count rates on the associated to the i-th channel and CC are the coincidence counts (see e.g. Ref. [46]). In our set-up we measured a channel efficiency, averaged over all 16 channels used, of -11.5 dB, which is an order of magnitude improvement respect to the experiment in chapter 3. From preliminary classical characterisation, such efficiency can be accounted to the different components as follows. The 16 SNSPDs presented an average efficiency of 0.78 (-1.0 dB). Mean transmission efficiencies through the off-chip filters and through the fibre connections to the detectors, averaged over the 16 channels used, were measured to be 0.87 (-0.57 dB) and 0.94 (-0.26 dB) respectively. The efficiency of the grating couplers

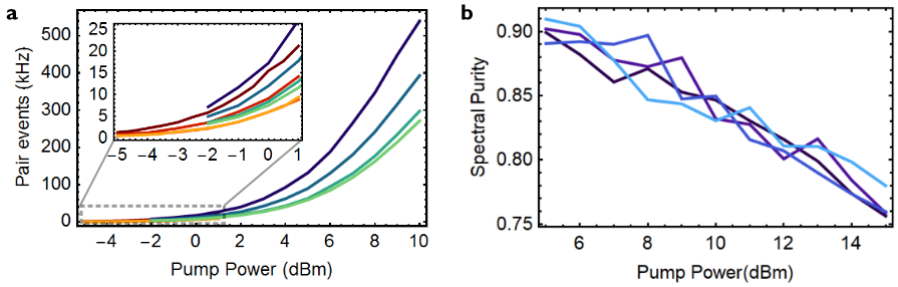


Figure 4.12: a, Photon pair (two-fold) event rates measured at the detectors. Blue-green lines are for non-degenerate SFWM while red-yellow lines are for degenerate SFWM. **b,** Purity of the signal photons emitted with non-degenerate SFWM, obtained via unheralded g_2 measurements.

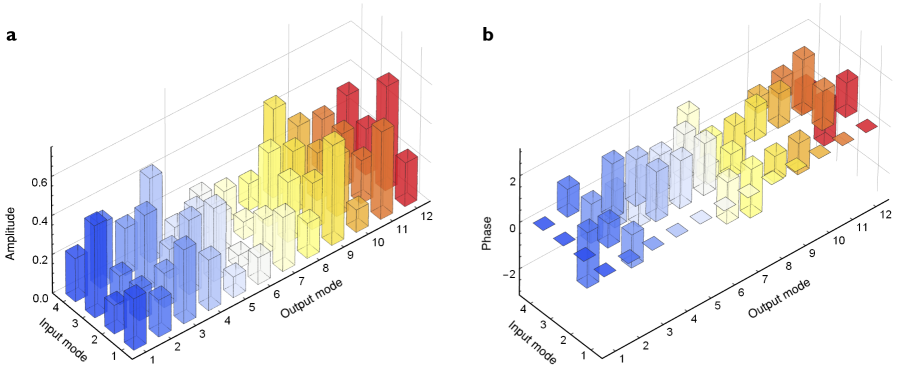


Figure 4.13: Amplitudes (**a**) and phases (**b**) of the measured entries of the 4×12 transfer matrix for the random walk.

is 0.77 (-1.1 dB) at the wavelengths used, and the transmission for the AMZI on-chip filters is 0.98 (-0.1 dB). Propagation loss through straight waveguides was estimated via cut-back measurements on test structures on the same silicon wafer to be approximately 2 dB cm^{-1}, indicating a transmission efficiency of 0.995 for the 12-mode random walk implemented. Due to bending losses, the propagation loss in the spiral waveguides is higher, with average losses of -7.1 dB measured in the 1.4 cm long spirals.

The performance of the individual sources was directly characterised by using the control grating couplers after the first array of AMZIs (outputs 1 in Fig 4.11), configuring the filters to collect the photons before the interferometer. The photon pairs are split off-chip[6] and detected by two SNSPDs. The measured two-fold coincidences count rates

6 The splitting of the photons is deterministic in the non-degenerate SFWM case, where idler and signal are separated using an off-chip filter. In the degenerate SFWM case a probabilistic splitting (with

from each single source, in both the degenerate and non-degenerate SFWM regimes, are reported in Fig. 4.12a as a function of the input (off-chip) pump power. To obtain these curves the two-fold coincidences from each source are analysed using the same two channels and detectors, with measured channel efficiencies of $\eta_1 = -9.7$ dB and $\eta_2 = -8.9$ dB.

In the non-degenerate SFWM regime, the coincidence count rates are given in term of the source efficiency (i.e. two-photon emission probability) $\epsilon = \tanh(\xi)^2/\cosh^2(\xi)$ as $CC = \epsilon R \eta_1 \eta_2$ [46, 47], where ξ is the squeezing parameter and R is the repetition rate of the laser, from which ϵ can be calculated. In the typical measurement conditions for scattershot boson sampling (9 dBm input pump power) we obtain for the four non-degenerate SFWM sources efficiencies of $\{6.5\%, 4.6\%, 3.4\%, 3.1\%\}$, that give the two-mode squeezing parameters $\boldsymbol{\xi} = \{0.25, 0.21, 0.18, 0.17\}$.

In the degenerate SFWM case, the characterisation of the sources can be performed similarly, with the only difference that, since the emitted photons are indistinguishable, the splitting now has to be performed probabilistically with a $50:50$ splitter. The coincidence count rates are thus now given by $CC = \epsilon R \eta_1 \eta_2/2$, where, generating single-mode squeezing, the source efficiency is $\epsilon = \tanh(\xi)^2/2\cosh(\xi)$. In Gaussian boson sampling measurement conditions (1 dBm total input pump power, measured off-chip) we obtain for the four degenerate SFWM sources efficiencies of $\{0.61\%, 0.41\%, 0.25\%, 0.27\%\}$, that give the single-mode squeezing parameters $\boldsymbol{\xi} = \{0.11, 0.09, 0.07, 0.07\}$.

A quantitative characterisation of the signal photons purity in the non-degenerate SFWM regime can be obtained via unheralded second order correlation ($g^{(2)}$) measurements [48]. To perform these measurements, photons are collected at the output of each source (through the AMZI filters typically used for pump filtering), the idler and signal photons are separated through an off-chip filter and second order correlation measurements are performed on the signal photon while the idler is not measured. Following Ref. [48], the photon purity $\mathcal{P}$ is estimated from the obtained second order correlation $g_2(0)$ via $\mathcal{P} = g_2(0) - 1$. In Fig. 4.12b the obtained purities for the four different sources are reported for various values of the input pump power. A decrease of the purity for increasing pump powers is observed, plausibly due to non-linear noises in the sources and multi-pair emission errors. In measurement conditions for scattershot boson sampling, the average purity is approximately 86%.

The transfer matrix of the implemented 4×12 continuous random walk was preliminarily characterised using standard methods based on classical and two-photon interference [49], and is pictured in Fig. 4.13.

probability 1/2) is performed via a fibre $50:50$ beam-splitter.

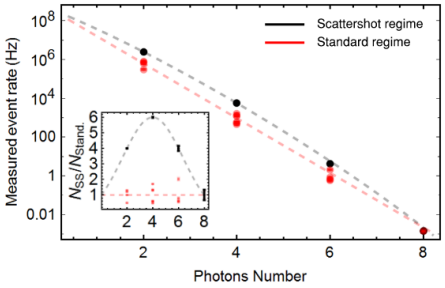


Figure 4.14: Measured event rates via non-degenerate SFWM for up to eight-fold photon coincidences in the scattershot (black data) and standard (red data) regimes. Inset: rate enhancement of the scattershot approach for different number of photons. Points are measured data and dashed lines are fits.

4.3 Results

In the experiment we focused on two type of results. From a technological perspective, we were interested in demonstrating integrated experiments with as many photons as possible generated and manipulated on a chip. In terms of applications, we were interested in implementing different boson-sampling experiments in the same device using the large number of generated photons, and switching between different boson sampling regimes using the pumping schemes described above.

4.3.1 *Generation rate of multi-photon states in the scattershot regime*

To generate photonic states in the scattershot regime we pumped the sources in the non-degenerate SFWM scheme, emitting weak two-mode squeezed states. In particular, when pumping all the 4 sources to generate $2n$-fold photon states, i.e. the generation of heralded states of n signal photons, a combinatorial improvement by $\binom{4}{n}$ is expected. Note that we can also implement the standard boson sampling approach by deciding to pump only n of the sources to generate n-photon states of the signals. Clearly, for $n = 4$ (8-fold coincidences) the two approaches are the same here. In Fig. 4.14 the rates directly measured at the detectors for $2n$-fold photon states are reported, both in the standard and scattershot boson sampling approaches. An average enhancement of approximately the desired combinatorial factor $\binom{4}{n}$ can be observed from the figure inset. Measured rates for four photons and six photons events, that is the generation of two and three heralded photons, are 5.8 kHz and 4 Hz, respectively. Eight photons coincidences are detected at a rate of approximately 4 events per hour.

4.3.2 *Performing scattershot boson sampling*

Boson sampling in the scattershot regime is obtained by delivering the generated states in the non-degenerate SFWM pumping into the random walk, and samples were obtained from the final photon counting. The measured input/output distributions p^{exp} in the

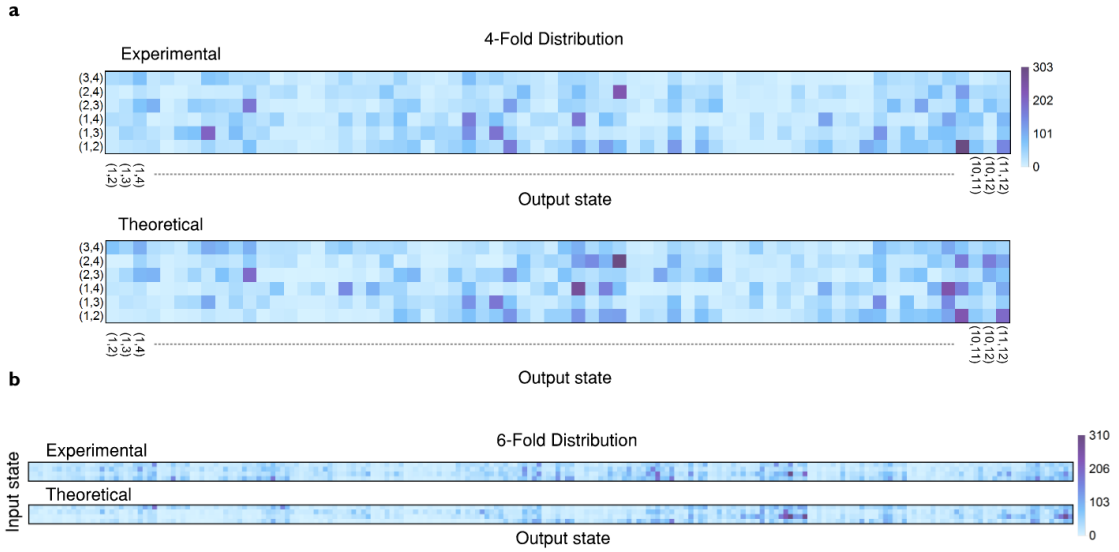


Figure 4.15: **a**, Experimental input/output distribution (top) of the four-fold events, showing 95% fidelity with the theoretical distribution (bottom). **b**, Experimental and theoretical input/output distribution of the six-fold events, with a fidelity of 92%. The collision-free output states on the horizontal axis are in increasing order $\{(1,2,3),(1,2,4),\ldots,(10,11,12)\}$ from left to right, and the input states on the vertical axis are $\{(1,2,3),(1,2,4),(1,3,4),(2,3,4)\}$ from bottom to top.

collision-free case are reported in Fig. 4.15a for the two heralded photons interference, obtained from over 14×10^3 4-fold events, and in Fig. 4.15b for three heralded photons interference, obtained from 27×10^3 6-fold events. We compare them with the theoretical distributions calculated via the permanent of a sub-matrix of the interferometer transmission matrix (see Fig. 4.3a). The 4-fold and 6-fold experimental distributions present a mean statistical fidelity

$$\mathcal{F}_s=\sum_i\sqrt{p_i^{\mathrm{exp}}p_i^{\mathrm{th}}} \tag{4.3.1}$$

of 95% and 92% with the theoretical distributions p^{th}, respectively.

Using the Bayesian validation methods, we validated the hypothesis of a scattershot boson sampler where the quantum interference is obtained from fully indistinguishable photons against a classical scenario where the photons are all fully distinguishable. In Fig. 4.16a and Fig. 4.16b the validation results are reported for 6-fold and 8-fold events, respectively. In both cases it is determined that samples are drawn from the quantum distribution with a confidence higher than 99.9%.

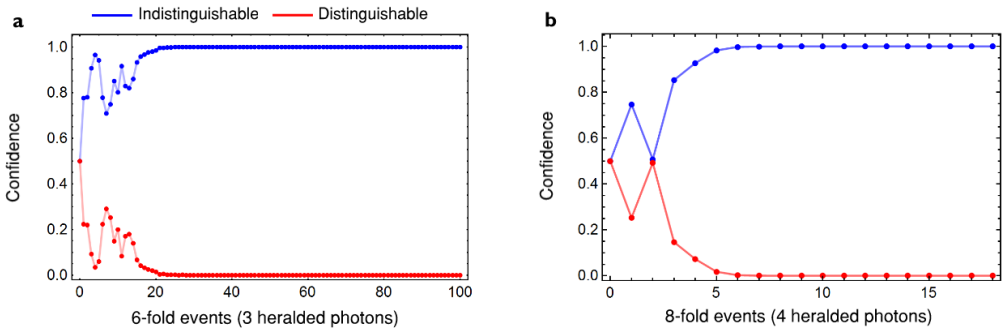


Figure 4.16: Validation of boson sampling from indistinguishable photons against classic sampling from fully distinguishable photons for SBS using six-photon (a) and eight-photon (b) data.

4.3.3 *Performing Gaussian boson sampling*

To implement Gaussian boson sampling we generate single-mode squeezed states via degenerate SFWM in the dual-pump scheme. Due to the lower pump power injected when using this scheme the event rates for Gaussian boson sampling are significantly lower than what we achieved for scattershot. We thus focused on events with detection of up to four signal photons, which we could observe at a rate of approximately a 5 Hz rate. The measured output four-photon distribution in the collision-free case is shown in Fig. 4.17. The mean statistical fidelity of the experimental output distribution with the ideal one, calculated using theoretical probabilities dictated by the Hafnian function, is 87%.

Gaussian boson sampling offers a wider range of scenarios to validate against compared to previous protocols. In fact, the theory of GBS allows one to calculate the output probability distribution for any physically relevant input Gaussian state. We therefore can validate against different noisy models which, to some extent, we expect to actually be present in our experiment. The validation results are shown in Fig. 4.18, while the details on how the output probabilities are measured for all tested models can be found in appendix D. First, we validate the ideal single-mode squeezing Gaussian boson sampling (SMSGBS) against the case where input thermal states are used (see Fig. 4.18a). Such states would be generated having two-mode squeezed states from the sources and without measuring one of the two modes, or from the presence of excessive losses (effectively analog to the multi-pair emission noise in standard boson sampling). Secondly, we validate against the case where the inputs are coherent states (see Fig. 4.18b), which may arise if the observed events would be dominated by contributions due to spurious pump photons. We go on to test the case where the input states are fully distinguishable single-mode squeezed states (see Fig. 4.18c), analog to the standard validation tests

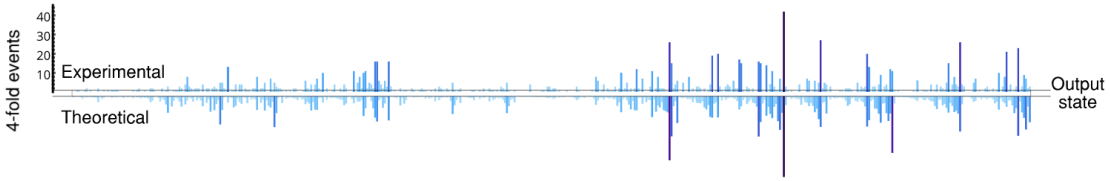


Figure 4.17: Experimental Gaussian boson sampling output distribution for the four-photon events, showing 87% fidelity with the theoretical distribution. The collision-free output states on the horizontal axis are in increasing order $\{(1,2,3,4),(1,2,3,5),\ldots,(9,10,11,12)\}$ from left to right.

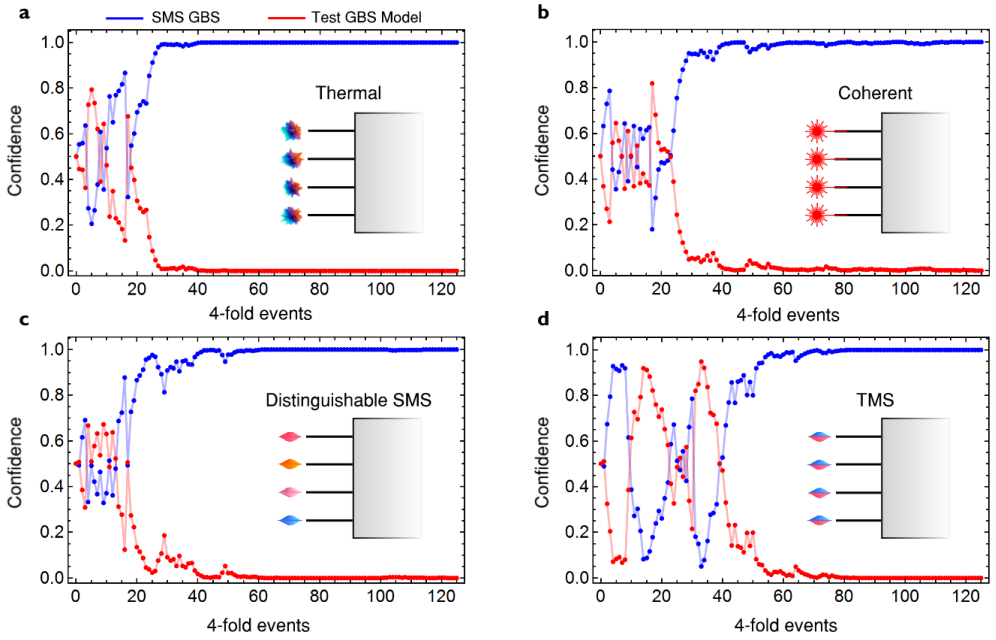


Figure 4.18: Validation of the ideal model of SMSGBS (blue lines) against different test models of imperfect implementations (red lines). The insets show a pictorial representation of the adversarial models: GBS with input thermal states (a), coherent states (b), distinguishable single-mode squeezed states (c), and two-mode squeezed (TMS) states (d). In all cases, the validation protocol is performed using Bayesian model comparison techniques.

in boson sampling. Namely, the input states are assumed correct but distinguishable, e.g. due to temporal or spectral mismatch. Finally we compare SMSGBS to the case where the input states are two-mode squeezed states (see Fig. 4.18d). This scenario can physically happen if we had photons generated by non-degenerate SFWM dominating the

observed events, instead of degenerate SFWM ones. For all tested models, a confidence higher than 99.9% in SMSGBS is obtained after approximately 120 samples.

4.4 Discussion

Passing from integrated experiments with the generation and manipulation of 2 and 4 photons, which were respectively first reported in 2014 [46] and few months before our result [2], to an 8 photons experiment is not a small technological step. For comparison, the first-ever 8-photon quantum state was generated only in 2012 [50] using bulk-optical techniques developed over decades of research activities. While we were finalising the paper, a new record was set in bulk-optics, demonstrating a 12-photon state [51]. Although integrated quantum photonics is a relatively much younger technology, our experiment gives reasons to be optimistic for its capabilities to reach and surpass such complexity in the near term. In particular, for the further scaling of the approach, some considerations can be drawn from this experience.

4.4.1 *Experimental considerations*

A significant challenge encountered in the experiment was setting up the detection system. In particular, achieving counting electronics fast enough to support scattershot or Gaussian boson sampling measurement regimes. To have some rough numbers in mind, we can calculate the average photon generation rate from each source via $\eta R\langle n\rangle$, where η is the loss per photon, R is the repetition rate of the pump and $\langle n\rangle$ is the mean number of photons generated at each pump pulse. Plugging in the values characterised, we obtain an event rate from each single source of approximately $6 \cdot 10^6$ events per second. Having four sources, we had to install counting electronics able to support more than 20 MTags/second distributed over 16 detector channels (note that, although this event rate is quite high, only few of such events happen simultaneously to provide an n-fold coincidence). Moreover, we had to develop and program a processor for the time-tags able to distil n-fold coincidences on the fly at the required speed, avoiding to save all the measured counts and doing that in post-processing. Increasing the target number of photons in such experiments would also increase (quadratically) the requirements in terms of electronic processing speed. For example, currently the highest speed commercial time-taggers support up to 500 MTags/second, which, using the system efficiencies of our experiment, would limit implementations to $\lesssim 100$ sources. We expect, however, that these numbers can be significantly improved with bespoke solutions. Additionally,

setting up the 16-detector system itself was not a simple task[7].

However, apart from these details on the detection electronics, which represents classical engineering challenges that, we believe, should be addressable, the results are in general very promising. The chip used to generate eight photons was itself basically plug-and-play; the control of the eight phase shifters in the AMZI filters was straightforward compared to the experiment in chapter 3, and the design, setting, and characterisation of the optical components in the chip is now a standard in silicon quantum photonics. The input/output chip coupling was stable for long periods of time and didn't represent a particular challenge. All these aspects indicate that, although a significant improvement has been shown in this experiments, we are still quite far from the ultimate scaling limits of the current technology.

4.4.2 *Rate estimations for near-term experiments*

It is then interesting to investigate what regimes of computational complexity we can aim at achieving with this approach, i.e. scaling up the component integration with the current silicon photonics technology. We can, for example, use the values measured in out set-up as relevant parameters to evaluate what event rates can be expected in near-term imperfect devices. The event rates for scattershot and Gaussian boson sampling experiments, in presence of losses and non-deterministic sources, and consisting of n signal photons emitted from k integrated sources and injected in a m mode interferometer, can be estimated via

$$R_{\mathtt{sbs}}(n,k,m) = R_0 \left[\binom{k}{n} \tanh(\xi)^{2n} \operatorname{sech}(\xi)^{2k} \right] \eta_{\mathtt{u}}^{mn} \eta_{\mathtt{ch}}^{2n} \eta_{\mathtt{det}}^{2n}, \tag{4.4.1}$$

$$R_{\mathtt{gbs}}(n,k,m) = R_0 \left[\binom{k/2+n/2-1}{n/2} \tanh(\xi)^{n} \operatorname{sech}(\xi)^{k} \right] \eta_{\mathtt{u}}^{mn} \eta_{\mathtt{ch}}^{n} \eta_{\mathtt{det}}^{n}. \tag{4.4.2}$$

Here, we have used R_0 as the repetition rate of the pump, ξ the squeezing of the sources, $\eta_{\mathtt{det}}$ the detection efficiency, and $\eta_{\mathtt{ch}}$ the system efficiency of the optical channels to the detectors (which includes chip-to-fiber coupling loss for off-chip detection). The term $\eta_{\mathtt{u}}^{m}$ indicates the losses inside the interferometer, where $\eta_{\mathtt{u}}$ represents the losses in each coupling operation, and the number of single operations each photon undergoes is m, as in the Clements scheme (see section 2.2.4). For simplicity, for both scatterhot and

7 The 16-detector system is shared between ≈ 10 experiments in our laboratory, and is located in a separate room accessible via a shared (quite lossy) fibre network. To avoid disrupting the rest of the laboratory, all the optical connections and measurements had to be performed during week-ends. In order to minimise losses, we avoided the shared fiber network by drilling a hole between the experiment room and the detector room, directly connecting the detectors to the chip set-up via low-loss short single-mode fibres.

# Signal Photons	Scattershot BS Event Rate (Hz)			
	WG	WG & Int.Det.	Rings	Rings & Int.Det.
10	0.3	2×10^3	65	4×10^5
16	-	8×10^{-3}	2×10^{-3}	3×10^3
20	-	-	-	30

Table 4.1: Estimated count rates for near-term scattershot experiments. The rates are calculated considering a silicon device where integrated sources deliver quantum states to a 100-modes interferometer ($m = k = 100$). The estimation is performed using the system efficiencies characterised in the experiment. Note that, to obtain the n signal photons experiments, $2n$ total photons need to be generated, including the n idler photons. *WG*: waveguide sources. *Rings*: sources based on micro-ring cavities. *Int.Det.*: cases where integrated detectors are considered.

# Signal Photons	Gaussian BS Event Rate (Hz)			
	WG	WG & Int.Det.	Rings	Rings & Int.Det.
10	6×10^3	5×10^5	6×10^4	5×10^6
16	1	1×10^3	4×10^2	4×10^5
20	2×10^{-3}	15	7	5×10^4
32	-	4×10^{-6}	1×10^{-5}	15
40	-	-	-	2×10^{-2}

Table 4.2: Estimated count rates for near-term Gaussian boson sampling experiments. The rates are calculated considering a silicon device where integrated sources deliver quantum states to a 100-modes interferometer ($m = k = 100$). The estimation is performed using the system efficiencies characterised in the experiment. *WG*: waveguide sources. *Rings*: sources based on micro-ring cavities. *Int.Det.*: cases where integrated detectors are considered.

Gaussian boson sampling, we will consider a source in each of the input modes of the interferometer ($k = m$). For the system efficiencies, we can use those characterised in our silicon photonic device and experimental set-up: $\eta_{\mathtt{det}} = 80\%$, $\eta_{\mathtt{ch}} = 64\%$, $\eta_{\mathtt{u}} = 99.95\%$, $R_0 = 500$ MHz.

In the analysis of the count rates for the silicon quantum photonics technology, we consider two standard types of integrated sources: a low efficiency one, i.e. the spiral sources used in our device, with a squeezing parameter $\xi = 0.17$, and more efficient sources based on integrated ring resonator cavities (see, for example, Ref. [47]), with $\xi = 0.31$. We also consider both the case where the detectors are off-chip, and the case where on-chip detection is performed. In the latter, no losses associated to off-chip

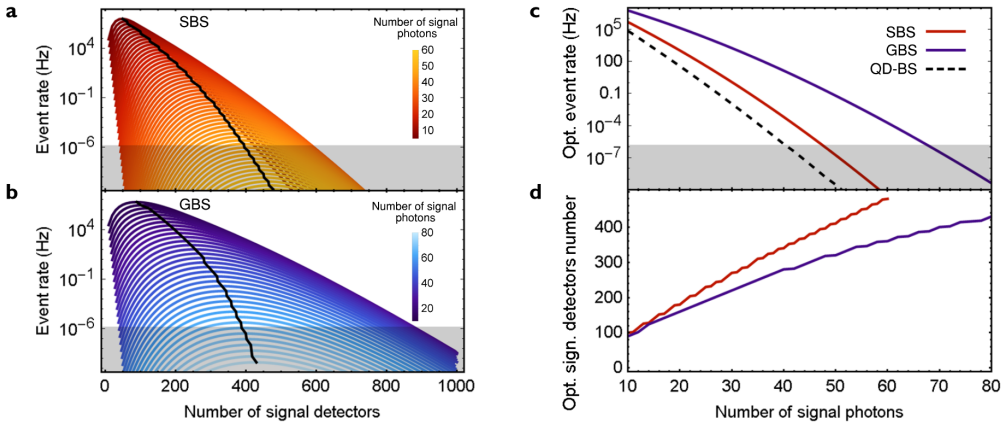


Figure 4.19: Estimated counts for different sizes of the optical circuit. In this analysis, the number of signal detectors is considered to be the same as the number of modes in the interferometer and the number of sources. **a** and **b** show the event rate for different circuit size and number of photons in a scattershot and Gaussian boson sampling scenario, respectively. Black lines show the optimal values of the event rates. **c,** Optimal event rate estimated for different numbers of signal photons. The dashed line represents an estimation for the case where standard boson sampling is performed using on-chip quantum dots sources [18]. **d,** Size of the interferometer associated to the optimal cases. In all plots, shaded areas represent impractical experiments, where the threshold is set to be 1 event/week.

coupling and fibre transmission are present ($\eta_{\text{ch}} = 1$).

The rates for these cases are reported in Table 4.1 and Table 4.2 for a scattershot and a Gaussian boson sampling scenario, respectively, in an interferometer with 100 modes. The results indicate that experiments with tens of photons are possible with current silicon photonics technologies. To further investigate the ultimate limits of the approach, we can test larger-scale circuits with up to 1000-modes interferometers and 1000 detectors on the signal modes, considering the most optimistic case of ring sources and integrated detectors. As can be observed in Fig. 4.19a-b, while increasing the number of sources provides an initial improvement to the event rates, it actually turns to be detrimental if the circuit size becomes too large. The reason for this is that losses in the interferometer become dominant for values of k that are too large, suppressing the combinatorial enhancement provided by the scattershot or Gaussian boson sampling approaches. Optimal event rates, reported in Fig. 4.19c, results from such trade-off, with the scaling of the size of the associated circuit shown in Fig. 4.19d. As expected, the presence of losses implies an exponential decrease in the event rate. However, experiments with a large number of photons are still possible before the event rate becomes impractical. In particular, Gaussian boson sampling can reach experiments with up to ≈ 70 signal photons, while the limit is ≈ 50 for scattershot boson sampling. These values are expected

to be at the limit of what is tractable with classical supercomputers [7].

To further increase the complexity, e.g. to target experiments with hundreds of photons, the capability to scale up the number of current optical components in silicon quantum photonics is not enough: technological progress is required. A first improvement would be to develop materials with lower transmission losses and more efficient sources of squeezed light. Promising alternative integrated photonic platforms are being developed in this direction [52]. Ultimately, the introduction of rudimentary error-correction techniques in boson sampling, tackling the effect of losses, is likely to be required for applications far beyond current capabilities.

Acknowledgements

The contributions of the author consisted in designing, mounting, and characterising the device and the experimental set-up, performing the experiment, and analysing the data. The practical organisation of the connections to the detector system and the measurement set up during weekends was done in cooperation with Ms. Caterina Vigliar and Dr. Raffaele Santagati. Dr. Jianwei Wang, Dr. Raffaele Santagati, Dr. Anthony Laing and Dr. Yunhong Ding provided support in the design of the experiment. The chip which was fabricated by Dr. Yunhong Ding at the Technical University of Denmark. Dr. Levon Chakhmakhchyan provided theoretical support in the data analysis and in investigating the test models for Gaussian boson sampling. The author acknowledges Mr. Joseph Lennon, Mr. Eric Johnston, Dr. Helmut Fedder and Dr. Michael Schlagmüller for really helpful assistance in programming the counting logic. We also thank Mr. Nicola Maraviglia, Ms. Rachel Chadwick, Dr. Chris Sparrow and Dr. Alex Neville for stimulating discussions on boson sampling. The work is published in Nature Physics (Ref. [1]).

References

1. Paesani, S., Ding, Y., Santagati, R., Chakhmakhchyan, L., Vigliar, C., Rottwitt, K., Oxenløwe, L. K., Wang, J., Thompson, M. G. & Laing, A. Generation and sampling of quantum states of light in a silicon chip. *Nature Physics* **15,** 925–929 (2019).
2. Faruque, I. I., Sinclair, G. F., Bonneau, D., Rarity, J. G. & Thompson, M. G. On-chip quantum interference with heralded photons from two independent micro-ring resonator sources in silicon photonics. *Optics express* **26,** 20379–20395 (2018).
3. Aaronson, S. & Arkhipov, A. The computational complexity of linear optics. *Quant. Inf. Comp.* **13,** 1383 (2014).
4. Bremner, M. J., Montanaro, A. & Shepherd, D. J. Achieving quantum supremacy with sparse and noisy commuting quantum computations. *Quantum* **1,** 8 (2017).
5. Trotzky, S., Chen, Y.-A., Flesch, A., McCulloch, I. P., Schollwöck, U., Eisert, J. & Bloch, I. Probing the relaxation towards equilibrium in an isolated strongly correlated one-dimensional Bose gas. *Nature Physics* **8,** 325 (2012).
6. Chakhmakhchyan, L. & Cerf, N. Boson sampling with Gaussian measurements. *Phys. Rev. A* **96,** 032326 (2017).
7. Neville, A., Sparrow, C., Clifford, R., Johnston, E., Birchall, P. M., Montanaro, A. & Laing, A. Classical boson sampling algorithms with superior performance to near-term experiments. *Nature Physics* **13,** 1153 (2017).
8. Dalzell, A. M., Harrow, A. W., Koh, D. E. & La Placa, R. L. How many qubits are needed for quantum computational supremacy? *arXiv:1805.05224* (2018).

9. Huh, J., Guerreschi, G. G., Peropadre, B., McClean, J. R. & Aspuru-Guzik, A. Boson sampling for molecular vibronic spectra. *Nature Photonics* **9,** 615 (2015).

10. Sparrow, C., Martín-López, E., Maraviglia, N., Neville, A., Harrold, C., Carolan, J., Joglekar, Y. N., Hashimoto, T., Matsuda, N., O'Brien, J. L., *et al.* Simulating the vibrational quantum dynamics of molecules using photonics. *Nature* **557,** 660 (2018).

11. González Olivares, D., Peropadre, B., Aspuru-Guzik, A. & García-Ripoll, J. J. Quantum simulation with a boson sampling circuit. *Phys. Rev. A* **94,** 022319. https://link.aps.org/doi/10.1103/PhysRevA.94.022319 (2016).

12. Crespi, A., Osellame, R., Ramponi, R., Brod, D. J., Galvao, E. F., Spagnolo, N., Vitelli, C., Maiorino, E., Mataloni, P. & Sciarrino, F. Integrated multimode interferometers with arbitrary designs for photonic boson sampling. *Nature Photonics* **7,** 545 (2013).

13. Tillmann, M., Dakić, B., Heilmann, R., Nolte, S., Szameit, A. & Walther, P. Experimental boson sampling. *Nature Photonics* **7,** 540 (2013).

14. Broome, M. A., Fedrizzi, A., Rahimi-Keshari, S., Dove, J., Aaronson, S., Ralph, T. C. & White, A. G. Photonic boson sampling in a tunable circuit. *Science* **339,** 794–798 (2013).

15. Spring, J. B., Metcalf, B. J., Humphreys, P. C., Kolthammer, W. S., Jin, X.-M., Barbieri, M., Datta, A., Thomas-Peter, N., Langford, N. K., Kundys, D., *et al.* Boson sampling on a photonic chip. *Science* **339,** 798–801 (2013).

16. Bentivegna, M., Spagnolo, N., Vitelli, C., Flamini, F., Viggianiello, N., Latmiral, L., Mataloni, P., Brod, D. J., Galvão, E. F., Crespi, A., *et al.* Experimental scattershot boson sampling. *Science Advances* **1,** e1400255 (2015).

17. Carolan, J., Harrold, C., Sparrow, C., Martín-López, E., Russell, N. J., Silverstone, J. W., Shadbolt, P. J., Matsuda, N., Oguma, M., Itoh, M., Marshall, G. D., Thompson, M. G., Matthews, J. C. F., Hashimoto, T., O'Brien, J. L. & Laing, A. Universal linear optics. *Science* **349,** 711–716 (2015).

18. Wang, H., He, Y., Li, Y.-H., Su, Z.-E., Li, B., Huang, H.-L., Ding, X., Chen, M.-C., Liu, C., Qin, J., *et al.* High-efficiency multiphoton boson sampling. *Nature Photonics* **11,** 361 (2017).

19. Loredo, J. C., Broome, M. A., Hilaire, P., Gazzano, O., Sagnes, I., Lemaitre, A., Almeida, M. P., Senellart, P. & White, A. G. Boson Sampling with Single-Photon Fock States from a Bright Solid-State Source. *Phys. Rev. Lett.* **118,** 130503. https://link.aps.org/doi/10.1103/PhysRevLett.118.130503 (2017).

20. Wang, H., Li, W., Jiang, X., He, Y.-M., Li, Y.-H., Ding, X., Chen, M.-C., Qin, J., Peng, C.-Z., Schneider, C., Kamp, M., Zhang, W.-J., Li, H., You, L.-X., Wang, Z., Dowling, J. P., Höfling, S., Lu, C.-Y. & Pan, J.-W. Toward Scalable Boson Sampling with Photon Loss. *Phys. Rev. Lett.* **120,** 230502. https://link.aps.org/doi/10.1103/PhysRevLett.120.230502 (2018).

21. Zhong, H.-S., Li, Y., Li, W., Peng, L.-C., Su, Z.-E., Hu, Y., He, Y.-M., Ding, X., Zhang, W.-J., Li, H., Zhang, L., Wang, Z., You, L.-X., Wang, X.-L., Jiang, X., Li, L., Chen, Y.-A., Liu, N.-L., Lu, C.-Y. & Pan, J.-W. 12-photon entanglement and scalable scattershot boson sampling with optimal entangled-photon pairs from parametric down-conversion. *arXiv:1810.04823* (2018).

22. Arkhipov, A. & Kuperberg, G. The bosonic birthday paradox. *Geometry & Topology Monographs* **18,** 1–7 (2012).

23. Renema, J., Shchesnovich, V. & Garcia-Patron, R. Quantum-to-classical transition in many-body bosonic interference. *arXiv:1809.01953* (2018).

24. Olson, J. P., Seshadreesan, K. P., Motes, K. R., Rohde, P. P. & Dowling, J. P. Sampling arbitrary photon-added or photon-subtracted squeezed states is in the same complexity class as boson sampling. *Phys. Rev. A* **91,** 022317. https://link.aps.org/doi/10.1103/PhysRevA.91.022317 (2015).

25. Seshadreesan, K. P., Olson, J. P., Motes, K. R., Rohde, P. P. & Dowling, J. P. Boson sampling with displaced single-photon Fock states versus single-photon-added coherent states: The quantum-classical divide and computational-complexity transitions in linear optics. *Phys. Rev. A* **91,** 022334. https://link.aps.org/doi/10.1103/PhysRevA.91.022334 (2015).

26. Rohde, P. P., Motes, K. R., Knott, P. A., Fitzsimons, J., Munro, W. J. & Dowling, J. P. Evidence for the conjecture that sampling generalized cat states with linear optics is hard. *Phys. Rev. A* **91,** 012342. https://link.aps.org/doi/10.1103/PhysRevA.91.012342 (2015).

27. Chabaud, U., Douce, T., Markham, D., van Loock, P., Kashefi, E. & Ferrini, G. Continuous-variable sampling from photon-added or photon-subtracted squeezed states. *Phys. Rev. A* **96,** 062307. https://link.aps.org/doi/10.1103/PhysRevA.96.062307 (2017).

28. Barkhofen, S., Bartley, T. J., Sansoni, L., Kruse, R., Hamilton, C. S., Jex, I. & Silberhorn, C. Driven Boson Sampling. *Phys. Rev. Lett.* **118,** 020502. https://link.aps.org/doi/10.1103/PhysRevLett.118.020502 (2017).

29. Wang, X.-J., Jing, B., Sun, P.-F., Yang, C.-W., Yu, Y., Tamma, V., Bao, X.-H. & Pan, J.-W. Experimental Time-Resolved Interference with Multiple Photons of Different Colors. *Phys. Rev. Lett.* **121,** 080501. https://link.aps.org/doi/10.1103/PhysRevLett.121.080501 (2018).

30. Lund, A. P., Laing, A., Rahimi-Keshari, S., Rudolph, T., O'Brien, J. L. & Ralph, T. C. Boson Sampling from a Gaussian State. *Phys. Rev. Lett.* **113,** 100502. https://link.aps.org/doi/10.1103/PhysRevLett.113.100502 (2014).

31. Hamilton, C. S., Kruse, R., Sansoni, L., Barkhofen, S., Silberhorn, C. & Jex, I. Gaussian Boson Sampling. *Phys. Rev. Lett.* **119,** 170501. https://link.aps.org/doi/10.1103/PhysRevLett.119.170501 (2017).

32. Pernice, W. H., Schuck, C, Minaeva, O, Li, M, Goltsman, G., Sergienko, A. & Tang, H. High-speed and high-efficiency travelling wave single-photon detectors embedded in nanophotonic circuits. *Nature Communications* **3,** 1325 (2012).

33. He, Y., Ding, X., Su, Z.-E., Huang, H.-L., Qin, J., Wang, C., Unsleber, S., Chen, C., Wang, H., He, Y.-M., Wang, X.-L., Zhang, W.-J., Chen, S.-J., Schneider, C., Kamp, M., You, L.-X., Wang, Z., Höfling, S., Lu, C.-Y. & Pan, J.-W. Time-Bin-Encoded Boson Sampling with a Single-Photon Device. *Phys. Rev. Lett.* **118,** 190501. https://link.aps.org/doi/10.1103/PhysRevLett.118.190501 (2017).

34. Kruse, R., Hamilton, C. S., Sansoni, L., Barkhofen, S., Silberhorn, C. & Jex, I. A detailed study of Gaussian Boson Sampling. *arXiv:1801.07488* (2018).

35. Arrazola, J. M. & Bromley, T. R. Using Gaussian Boson Sampling to Find Dense Subgraphs. *Phys. Rev. Lett.* **121,** 030503. https://link.aps.org/doi/10.1103/PhysRevLett.121.030503 (2018).

36. Caianiello, E. R. *Combinatorics and Renormalization in Quantum Field Theory* (W. A. Benjamin, Inc., 1973).

37. Björklund, A. *Counting perfect matchings as fast as Ryser* in *Proceedings of the twenty-third annual ACM-SIAM symposium on Discrete Algorithms* (2012), 914–921.

38. Aaronson, S. & Arkhipov, A. BosonSampling Is Far From Uniform. *Quantum Information and Computation* **14,** 1383–1423 (2014).

39. Bentivegna, M., Spagnolo, N., Vitelli, C., Brod, D. J., Crespi, A., Flamini, F., Ramponi, R., Mataloni, P., Osellame, R., Galvão, E. F., *et al.* Bayesian approach to boson sampling validation. *International Journal of Quantum Information* **12,** 1560028.

40. Spagnolo, N., Vitelli, C., Bentivegna, M., Brod, D. J., Crespi, A., Flamini, F., Giacomini, S., Milani, G., Ramponi, R., Mataloni, P., *et al.* Experimental validation of photonic boson sampling. *Nature Photonics* **8,** 615 (2014).

41. Ding, Y, Ou, H & Peucheret, C. Ultra-high-efficiency apodized grating coupler using fully etched photonic crystals. *Opt. Lett.* **38,** 2732–2734 (2013).

42. Ding, Y., Ou, H. & Peucheret, C. Ultrahigh-efficiency apodized grating coupler using fully etched photonic crystals. *Opt. Lett.* **38,** 2732–2734 (2013).

43. Ding, Y., Peucheret, C., Ou, H. & Yvind, K. Fully etched apodized grating coupler on the SOI platform with -0.58 dB coupling efficiency. *Opt. Lett.* **39,** 5348–5350 (2014).

44. Banchi, L., Burgarth, D. & Kastoryano, M. J. Driven Quantum Dynamics: Will It Blend? *Phys. Rev. X* **7,** 041015. https://link.aps.org/doi/10.1103/PhysRevX.7.041015 (2017).

45. Muraleedharan, G., Miyake, A. & Deutsch, I. H. Quantum computational supremacy in the sampling of bosonic random walkers on a one-dimensional lattice. *arXiv:1805.01858* (2018).

46. Silverstone, J. W., Bonneau, D., Ohira, K., Suzuki, N., Yoshida, H., Iizuka, N., Ezaki, M., Natarajan, C. M., Tanner, M. G., Hadfield, R. H., *et al.* On-chip quantum interference between silicon photon-pair sources. *Nature Photonics* **8,** 104 (2014).

47. Silverstone, J. W., Santagati, R., Bonneau, D., Strain, M. J., Sorel, M., O'Brien, J. L. & Thompson, M. G. Qubit entanglement between ring-resonator photon-pair sources on a silicon chip. *Nature Communications* **6,** 7948 (2015).

48. Christ, A., Laiho, K., Eckstein, A., Cassemiro, K. N. & Silberhorn, C. Probing multimode squeezing with correlation functions. *New J. Phys.* **13,** 033027 (2011).

49. Laing, A. & O'Brien, J. L. Super-stable tomography of any linear optical device. *arXiv:1208.2868v1* (2012).

50. Yao, X.-C., Wang, T.-X., Xu, P., Lu, H., Pan, G.-S., Bao, X.-H., Peng, C.-Z., Lu, C.-Y., Chen, Y.-A. & Pan, J.-W. Observation of eight-photon entanglement. *Nature photonics* **6,** 225–228 (2012).

51. Zhong, H.-S., Li, Y., Li, W., Peng, L.-C., Su, Z.-E., Hu, Y., He, Y.-M., Ding, X., Zhang, W.-J., Li, H., *et al.* 12-photon entanglement and scalable scattershot boson sampling with optimal entangled-photon pairs from parametric down-conversion. *arXiv:1810.04823* (2018).

52. Dutt, A., Luke, K., Manipatruni, S., Gaeta, A. L., Nussenzveig, P. & Lipson, M. On-Chip Optical Squeezing. *Phys. Rev. Applied* **3,** 044005. https://link.aps.org/doi/10.1103/PhysRevApplied.3.044005 (2015).

5 ANALOG QUANTUM SIMULATION WITH PHOTONS

After having described the technological advances toward large-scale integrated quantum photonics experiments, in the next few chapters we will discuss some possible future applications for this platform. In particular, we will focus on quantum simulation applications, which will be presented in increasing order of complexity in terms of requirements for the quantum hardware.

The first one in this order is the photonic analog quantum simulation of molecular properties. The idea behind this type of application is to map vibrations in molecules into a boson sampler, with the output statistics being linked to the desired property of a target molecule. We will discuss two protocols of this type: the simulation of vibrational molecular dynamics [2], and the calculation of molecular vibronic transition probabilities [3] (also known as Franck-Condon coefficients). For the latter, we will report an experimental investigation to test its performance on the silicon device already reported in chapter 4.

Apart from the applications discussed here, boson sampling has been mapped also to other types of problems, including the simulation of spin Hamiltonians [4], the enhancement of classical optimisation heuristics [5], and graph theory calculations [6, 7]. Although it would be interesting to test such applications in near-term devices to investigate their practicality, they are not discussed in this thesis.

5.1 Molecular quantum dynamics simulation via boson sampling

The reason why complex quantum chemical systems are intractable on classical machines was described by Dirac already in the early days of quantum mechanics [8]. Namely, he noted that the wavefunction of a quantum system grows exponentially with the number

of particles, making classical computers unable to exactly simulate quantum systems in an efficient way. This problem led Feynman in 1982 to propose the development of controllable quantum hardware for efficient simulation of complex chemical systems [9]. Quantum chemistry can thus be considered as the original motivation that led to the field of quantum computing in the first place.

In this chapter the particular quantum chemistry problems we address are vibrational dynamics of molecules, and transitions between different vibrational states induced by an electronic process in the molecule. We will describe, quite informally, the framework used to describe these problems, with a focus on mapping them to evolutions in quantum photonics. The general underlying idea is to exploit the same nature of bosonic excitations in quantised vibrational modes (phonons) and bosonic excitations of the electromagnetic field (photons).

5.1.1 *Molecular vibrations in the harmonic approximation*

Vibrational dynamics in a molecule are essentially small oscillations of the nuclei around a local minimum in the potential energy surface of the molecule. The energy surface depends on the electronic structure of the molecule and the nuclear positions. In the Born-Oppenheimer approximation, the electronic and nuclear degrees of freedom are decoupled, so that the potential energy surface can be considered independent of the vibrational state of the molecule [10]. For a general molecule with N atoms the energy surface lives in a $3N-6$ space, so that $3N-6$ vibrational modes are possible.[1] As in classical mechanics, for treating such oscillations it is very convenient to approximate the potential energy near the minimum as a quadratic form, as shown in Fig. 5.1. This is the so called *harmonic approximation*, valid in the limit of small vibrational energies. Under this approximation all vibrations can be treated as harmonic oscillators. The $3N-6$ normal vibrational modes and the associated vibrational frequencies $\{\omega_i\}$ can be calculated by diagonalising the Hessian matrix of the potential energy surface in the minimum, equivalently as in the classical treatment of oscillations. The calculation of the potential energy surface and the eigenspectrum of the Hessian are central concepts in computational chemistry [11], and standard techniques from quantum chemistry can

1 The number of vibrational degrees of freedom can be counted in the following way. While $3N$ are the total Cartesian coordinates required to describe all atoms, the position of the centre of mass and the angles relative to it can be chosen arbitrarily, and thus represent no real degrees of freedom. In other words, translations or rotations of the whole systems do not represent vibrations but only global transformations. We therefore have to subtract $3+3$ degrees of freedom from the total $3N$, obtaining $3N-6$. If the molecule is linear, there are only two rotational degrees of freedom for the whole molecule, as the rotation around the axis of the molecule is excluded. In this case the total degrees of freedom are then $3N-5$.

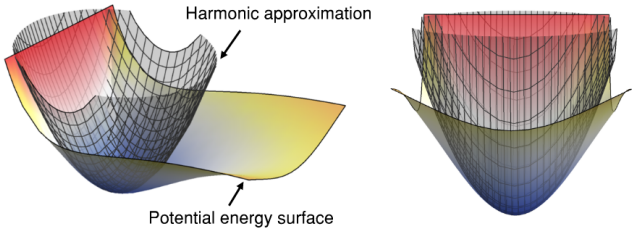


Figure 5.1: Example of an anharmonic potential energy surface. The harmonic approximation is obtained by approximating the function near a local minimum with a quadratic form.

provide approximate practical calculations[2] for molecules with hundreds of vibrational modes [12]. In a quantum mechanical description, the dynamics of the independent harmonic oscillators, i.e. the normal modes, are defined via the Hamiltonian

$$\hat{H} = \sum_{\ell=1}^{3N-6} \hbar\omega_\ell \hat{a}_\ell^\dagger \hat{a}_\ell, \tag{5.1.1}$$

where $\hat{a}_\ell^\dagger$ and $\hat{a}_\ell$ are the bosonic creation and annihilation operators of the ℓ-th normal vibrational mode, with associated frequency ω_ℓ.

Apart from the normal modes, a type of vibrational modes that are of practical interest are the so called *localised modes*. These are modes where the vibrational energy is spatially localised on single atoms of the molecule, as shown in Fig. 5.2; they are found by numerically maximising the sum of the squares of kinetic energy contributions from each nucleus. Localised modes are of practical importance for understanding many molecular phenomena, such as energy transport and dissociation, and single excitations of these modes can be prepared in quantum chemistry experiments [13]. We will therefore focus on the quantum dynamics of molecules prepared and measured in localised modes.

In contrast to normal modes, localised modes are not eigenstates of the Hamiltonian in eq.(5.1.1), therefore coherent population transfers between the different oscillatory modes happen during the evolution. In order to describe such dynamics it is convenient to define the basis transformation between the normal and localised modes, which is given by a unitary matrix U^L via

$$\hat{a}_\ell^\dagger \mapsto \sum_{k=1}^{3N-6} U^L_{k,\ell} \hat{a}_k^\dagger. \tag{5.1.2}$$

2 Classical methods to calculate potential energy surfaces, which depends on the electronic structure of the molecule, are typically only reliable for molecules where the electronic wavefunction presents weak correlations. For molecular systems with stronger correlations between the electrons, the approximations used to make classical approaches practical cease to be reliable, and the energy calculation generally becomes intractable on classical machines. Luckily, in some cases quantum computers could offer efficient solutions, as is discussed in chapter 6.

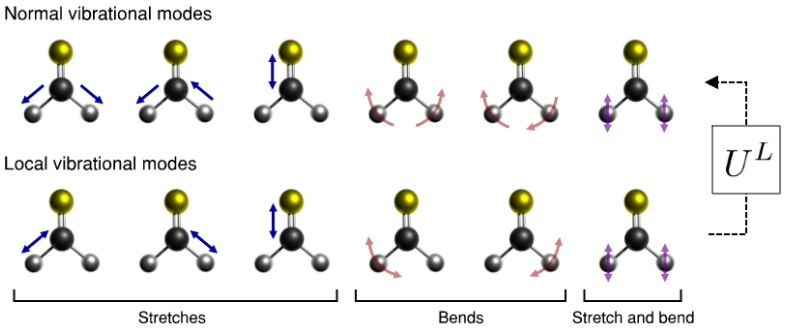


Figure 5.2: Representation of the normal (top) and localised (bottom) modes for vibrations in the H_2CS (Thioformaldehyde) molecule. The change of basis between the localised and normal modes is performed via the unitary matrix U^L.

The unitary matrix U^L thus represents a simple change of basis, and can be readily calculated once the normal and localised modes are known. At this point, a description in the localised basis of the unitary dynamics can be obtained following a simple prescription (pictured in Fig. 5.3a):

1. Apply the unitary U^L to convert from the localised modes basis to the normal modes basis.
2. Independently evolve the normal modes for a time t according to the Hamiltonian in eq.(5.1.1), which provides the unitary evolution $\bigoplus_\ell e^{-iw_\ell t/\hbar}$.
3. Apply the unitary $U^{L\dagger}$ to convert back from the normal modes basis to the localised modes basis.

More explicitly, the evolution of the localised modes at time t is given by the unitary:

$$U(t) = U^{L\dagger}\left(\bigoplus_\ell e^{-iw_\ell t/\hbar}\right)U^L. \tag{5.1.3}$$

5.1.2 *Mapping to boson sampling*

It is now straightforward to see the connection between the dynamics of localised vibrational modes and boson sampling. Vibrational excitations correspond to the creation of bosons in the associated modes, and the dynamics are described by a unitary operation $U(t)$. Each of these ingredients can be exactly mapped into a photonic scenario in the following way:

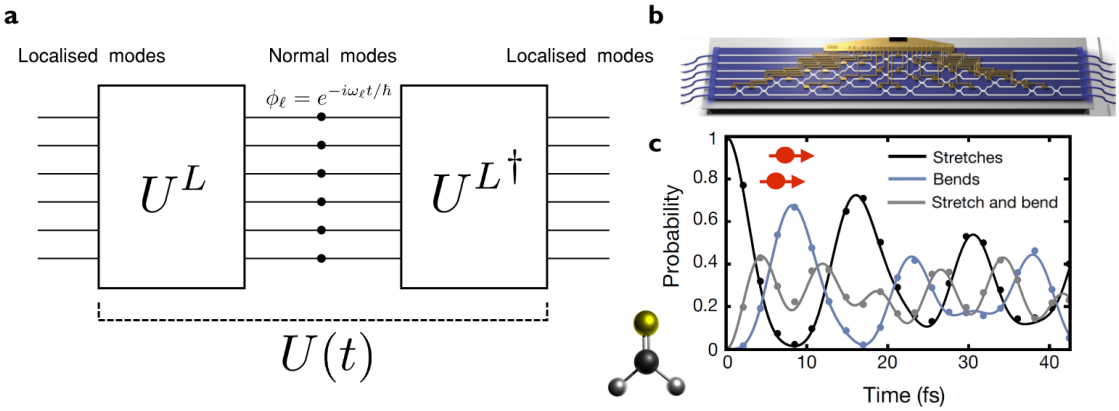


Figure 5.3: a), The quantum vibrational dynamics of the molecule can be described in the localised basis by converting to the normal basis via U^L, evolution in the local basis for a time t, and then convert back in the localised basis via $U^{L\dagger}$. **b)**, Silica device used for the first experimental demonstration of the simulation algorithm. Image taken from Ref. [14]. **c)**, Example of experimentally simulated quantum dynamics in the H_2CS (Thioformaldehyde) molecule, reported in Ref. [2].

- Vibrational modes $\leftrightarrow$ Optical modes.
- Vibrational mode ℓ initialised with n excitations $\leftrightarrow$ Optical mode ℓ initialised with n photons.
- Evolution of the molecular vibrations described by $U(t) \leftrightarrow$ Evolution of the photons in an interferometer described by $U(t)$.
- Measurement of the final vibrational configuration $\leftrightarrow$ Photon-number detection at the output of the interferometer.

The experimental scenario described above for the simulation of vibrational dynamics is equivalent to the boson sampling schemes discussed in chapter 4: photons need to be prepared in an input state that matches the initial vibrational state of the molecule, and output configurations are sampled after the evolution according to $U(t)$. In particular, simulating molecules that are initially prepared in Fock vibrational states, i.e. with a fixed number of excitations, corresponds to the standard boson sampling protocol. Note also that one could think of preparing the molecules in squeezed states (or a subset of modes in other Gaussian states, e.g. coherent states or thermal states); the quantum dynamics in this case are mapped into Gaussian boson sampling.

Therefore, the simulation of molecular quantum dynamics in the harmonic regime

essentially inherits both the efficiency of boson sampling protocols[3] (with an additional computational overhead for calculating the normal and localised modes, which is typically polynomial [11, 12]) and experimental challenges to implement it on a large scale.

The first demonstration of this quantum simulation protocol was performed by C. Sparrow et al. at the University of Bristol [2], where initial Fock photonic states (up to 4 photons) were generated off-chip and injected into a silica chip (shown in Fig. 5.3b) where a reconfigurable universal 6-mode interferometer acted as a re-programmable molecule. Exploiting this capability to implement arbitrary evolutions $U(t)$, they simulated the quantum dynamics for a wide variety of molecules (limited to 6 vibrational modes). An example from their work is reported in Fig. 5.3c where the statistics of the vibrational state of the H_2CS molecule, initially prepared in a superposition of two excitations, are shown for various values of the evolution time t.

As with boson sampling, experiments involving $\gtrsim 20$ are expected to enter an interesting regime where simulation problems practically hard for classical approaches will start to be accessible. Technological developments as those reported in chapters 3 and 4 provide a promising route to reach such regimes. We are currently working toward this goal. In the meanwhile, it is important to test these analog quantum simulation algorithms on current quantum hardware to investigate if the technology can actually cope with the requirements of these applications. In the next section such a test is performed using a different analog photonic quantum chemistry simulation task.

5.2 Calculation of Franck-Condon profiles with photonics

A different quantum chemistry problem that can be mapped to boson sampling is the calculation of molecular vibronic (vibrational and electronic) spectra, known as Franck–Condon (FC) profiles. In the previous section we focused on studying the dynamics of molecular oscillations; here we are instead interested in transformations of the oscillatory modes when the electrical structure of the molecule is perturbed. Vibronic spectroscopy is an important field in quantum chemistry, as the spectra are important tools to investigate molecular properties. While absorption spectra of molecules determine

3 While its appear quite plausible that the computational complexity of the simulation of molecular quantum dynamics is approximately the same as boson sampling, due to the mapping just described, such connection is not necessarily correct in general. In fact, some of the technical assumptions required to formally demonstrate the complexity of boson sampling are difficult to apply in a simulation scenario. For example, is the distribution of the dynamics $U(t)$ Haar random? Probably no. Therefore, in general we need to be careful when asserting the computational complexity of molecular simulations. On the other hand, in contrast to boson sampling, we here are clearly not concerned in testing axioms of the theory of computation. Rather, the spirit is to try to solve problems that are intractable with known quantum chemistry approaches, notwithstanding the underlying classical computational complexity.

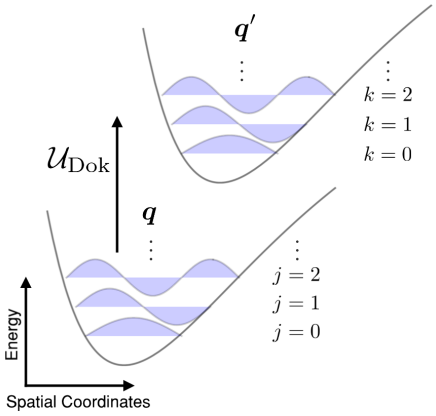


Figure 5.4: Simplified schematic of a vibronic transition. The molecule is initially in an electronic configuration with normal vibrational modes (in the harmonic approximation) $\boldsymbol{q}$ associated to the energy surface. When a process, e.g. photon absorption, induces a change in the electronic structure, the energy surface is modified, defining a new set of normal vibrational modes $\boldsymbol{q}'$. In the Franck-Condon approximation, the transformation between the two sets of vibrational modes is given by a linear mapping $\mathcal{U}_{\text{Dok}}$.

industrially relevant properties, such as the performance as solar cells [15] or as dyes [16], their prediction using classical approaches is computationally hard already for molecules of modest size [17–19]. A more efficient way to calculate Franck-Condon profiles would thus be a potentially groundbreaking achievement, both scientifically and industrially.

In 2015 Hu et al., in Aspuru-Guzik's quantum chemistry group at Harvard, have formulated an interesting approach to calculate Franck-Condon profiles with machines similar to Gaussian boson samplers [3], potentially providing significant speed-ups. We will first describe theory of the approach, and then report an experimental implementation of the protocol on a silicon photonic chip.

5.2.1 *Vibronic transition spectra*

As discussed in the previous section, molecular vibrations represent small oscillations in the potential energy surface. In particular, under the harmonic approximation, they are described by a set of independent quantum harmonic oscillators. In such a description, the potential energy surface is considered constant it time, i.e. we assume the atoms are in a steady electronic configuration. However, electronic transitions can happen, for instance, due to molecular processes such as photon absorption or emission. When an electronic transition happens, the potential energy surface is modified, inducing a new set of vibrational modes. Vibronic transition profiles represent the transition probabilities between an initial vibrational mode $|\boldsymbol{j}\rangle$ to a final one $|\boldsymbol{k}\rangle$ after an electronic transition has happened. For simplicity, we consider the molecule at a temperature of ≈ 0 K, in which the initial state contains no vibrations[4], i.e. $|\boldsymbol{j}\rangle = |0, \ldots, 0\rangle$, while the final state

4 In general, for a non-zero temperature $T > 0$, the initial state is a thermal state, i.e. a mixture of the vibrational states with each term weighted by $\exp(-\frac{E}{kT})$, where E is the energy associated to the vibrational state [3]. As thermal states are still Gaussian states, the analogy between the quantum

$|\boldsymbol{k}\rangle = |k_1, \ldots, k_m\rangle$ can in general have a non-zero number of vibrations k_i in the i-th vibrational mode. Here $i \in \{1, \ldots, m\}$ with $m = 3N - 6$ the total number of vibrational modes ($m = 3N - 5$ for a linear molecule).

In the formalism used to describe vibronic transitions, first proposed by Duschinsky [20], the transformation of the vibrational normal coordinates $\boldsymbol{q}$ of the molecule is characterised via a linear relation

$$q_i \to \sum_{j=1}^{m} U_{i\ell} q_\ell + d_i, \tag{5.2.1}$$

where U, known as the Duschinsky rotation, is a orthogonal matrix (unitary matrix with real entries), and $\boldsymbol{d}$ a displacement vector (again with real entries). In reconstructing the transition spectra, U and $\boldsymbol{d}$ characterised the particular process undergone by the molecule, and thus represent, together with the initial and final frequencies of the normal vibrational modes ($\{\omega_i\}$ and $\{\omega'_i\}$ respectively), the input parameters to the problem.

According to the Franck-Condon principle [21, 22], the probability of a vibronic transition between $|\boldsymbol{j}\rangle$ and $|\boldsymbol{k}\rangle$ is given by

$$p_{\mathrm{FC}}(\boldsymbol{k}) = |\langle \boldsymbol{k}|\mathcal{U}_{\mathrm{Dok}}|0, \ldots, 0\rangle|^2. \tag{5.2.2}$$

Here, $\mathcal{U}_{\mathrm{Dok}}$ is known as the Doktorov transformation [23] and, following the theory in Ref. [3], can be conveniently written as

$$\mathcal{U}_{\mathrm{Dok}} = \mathcal{U}_{\mathrm{L}} \left[\otimes_{i=1}^{m} \mathcal{S}(\xi_i)\right] \mathcal{U}_{\mathrm{R}}^{\dagger} \left[\otimes_{i=1}^{m} \mathcal{D}(\alpha_i)\right] \tag{5.2.3}$$

which is a combination of single-mode displacements

$$\mathcal{D}(\alpha_i) = \exp\left(\alpha_i \hat{a}_i + \alpha_i^* \hat{a}_i^{\dagger}\right), \tag{5.2.4}$$

unitary evolutions $\mathcal{U}_{\mathrm{L}}$ and $\mathcal{U}_{\mathrm{R}}$ (which conserve the number of excitations), and single-mode squeezing operations

$$\mathcal{S}(\xi_i) = \exp\left[\xi_i/2(a_i^2 + a_i^{\dagger 2})\right]. \tag{5.2.5}$$

The displacement and squeezing parameters on the i-th mode are indicated via α_i and ξ_i respectively, and $\hat{a}_i^{\dagger}$ ($\hat{a}_i$) are the bosonic creation (annihilation) operators associated to the vibrations in the molecule (i.e. phonons).

The parameters in eq.(5.2.3), that is the unitary matrices U_{L} and U_{R} associated to the transformations $\mathcal{U}_{\mathrm{L}}$ and $\mathcal{U}_{\mathrm{R}}$, and the values of $\boldsymbol{\alpha}$ and $\boldsymbol{\xi}$, can be calculated from the problem inputs (U,$\boldsymbol{d}$, $\{\omega_i\}$, and $\{\omega'_i\}$) as follows [3]. Defining the matrices

simulation protocol and Gaussian boson sampling holds also for simulations at non-zero temperatures. A thermal state at $T = 0$K represents the vacuum (see section 2.1), i.e. no vibrations are present.

$$J = \Omega' U \Omega^{-1}, \qquad \Omega = \mathrm{diag}(\sqrt{\omega_1}, \ldots, \sqrt{\omega_m}), \qquad \Omega' = \mathrm{diag}(\sqrt{\omega'_1}, \ldots, \sqrt{\omega'_m}), \tag{5.2.6}$$

the displacement parameters are obtained via

$$\alpha_i = \frac{1}{\sqrt{2}} \sum_{j=1}^{m} J^{-1}_{ij} \delta_j, \tag{5.2.7}$$

with $\delta_i = \sum_{j=1}^{m} \Omega'_{ij} d_j$. The unitaries U_{L} and U_{R} can be found via the singular value decomposition of the matrix J, with the squeezing parameters related to the singular eigenvalues via

$$J = U_{\mathrm{L}} \Sigma U_{\mathrm{R}}^{\top}, \qquad \xi_i = \ln(\Sigma_{ii}). \tag{5.2.8}$$

Finally, the molecular Franck-Condon profile at a given frequency ω reads

$$\mathrm{FCP}(\omega) = \sum_{k_1,\ldots,k_m=0}^{\infty} p_{\mathrm{FC}}(\boldsymbol{k}) \delta \left(\omega - \sum_{i=1}^{m} \omega'_i k_i \right), \tag{5.2.9}$$

where $\delta(x)$ is the Kronecker delta function and $p_{\mathrm{FC}}(\boldsymbol{k})$ is the transition probability for the final vibrational configuration $\boldsymbol{k}$ defined in eq.(5.2.2). Note that, since $p_{\mathrm{FC}}(\boldsymbol{k})$ is a probability distribution over the final configurations $\boldsymbol{k}$, the Franck-Condon profile formally represents the average value of the delta function $\delta(\omega - \sum_{i=1}^{m} \omega'_i k_i)$.

An important problem for classical approaches is the estimation of the Franck-Condon factors $p_{\mathrm{FC}}(\boldsymbol{k})$ [3]. Although no formal arguments of the computational complexity of such task exist, the relation with a generalised boson sampling problem provides an intuition of the intractability of the calculation of Franck-Condon profiles on classical machines.

5.2.2 *Mapping to Gaussian boson sampling*

The mapping of Franck-Condon profiles to a boson sampling protocol is performed by exploiting the formulation of the Doktorov transformation in terms of Gaussian operations, as in eq. (5.2.3). Combining it with eq. (5.2.2), we can rewrite the Franck-Condon factors as

$$p_{\mathrm{FC}}(\boldsymbol{k}) = \left| \langle \boldsymbol{k} | \mathcal{U}_{\mathrm{L}} \left[\otimes_{i=1}^{m} \mathcal{S}(\xi_i) \right] \mathcal{U}_{\mathrm{R}}^{\dagger} \left[\otimes_{i=1}^{m} \mathcal{D}(\alpha_i) \right] |0, \ldots, 0\rangle \right|^2 \tag{5.2.10}$$

$$= \left| \langle \boldsymbol{k} | \mathcal{U}_{\mathrm{L}} | \psi \rangle \right|^2, \tag{5.2.11}$$

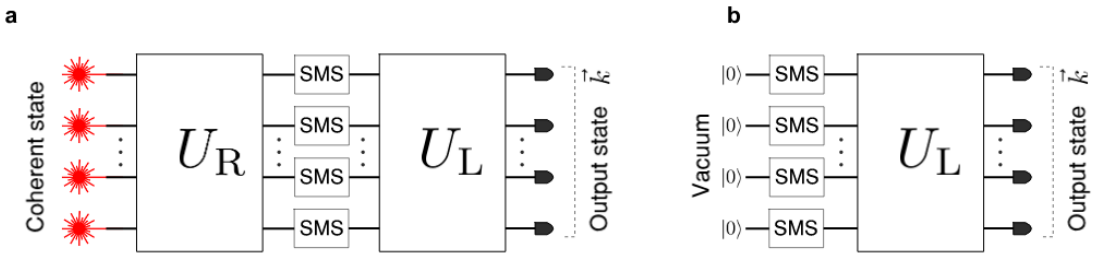


Figure 5.5: a), Schematic of the photonic circuit required for the calculation of Franck-Condon profiles. **b)**, Simplified schematic in the zero-displacement case, which is now equivalent to the Gaussian boson sampling scenario reported in chapter 4.

where the state $|\psi\rangle$ is defined by

$$|\psi\rangle := \left[\otimes_{j=1}^{m} \mathcal{S}(\xi_j)\right] \mathcal{U}_{\mathrm{R}}^{\dagger} \left[\otimes_{i=1}^{m} \mathcal{D}(\alpha_i)\right] |0, \ldots, 0\rangle \,. \tag{5.2.12}$$

It is easy to notice that $|\psi\rangle$ is a squeezed-coherent state, that is a Gaussian state with covariance $\sigma_0 = \frac{1}{2} S(\boldsymbol{\xi}) S(\boldsymbol{\xi})^{\dagger}$ (with $S(\boldsymbol{\xi})$ the squeezing symplectic matrix, see eq. 2.1.37) and displacement $U_{\mathrm{R}}^{\top} \boldsymbol{\alpha}$, as pictured in Fig. 5.5a. The scenario described in eq.(5.2.11) is then very similar to the general Gaussian boson sampling protocol described in chapter 4 (see Fig. 4.6a): a Gaussian input state $|\psi\rangle$ is prepared, evolved according to a unitary transformation U_{L}, and measured at the output with photon counters to estimate the probability $p_{\mathrm{FC}}(\boldsymbol{k})$.

While the mapping between the molecular parameters and photonics is similar to the one described for the mapping of molecular dynamics in section 5.1 (vibrational modes and excitations are mapped to optical modes and photons), important differences are present. In general, the main different features with the protocols for molecular dynamics simulation and boson sampling algortihms are:

- Input state. In general we need displaced squeezed states, rather than input Fock states (as in standard or scattershot boson sampling) or Gaussian states with zero displacement (as in Gaussian boson sampling with single mode squeezing).

- Interpretation of the operations. The initial state and the unitary operation implemented are now intimately connected and provide altogether the desired transformation, but no longer have a physical meaning on their own. This is in contrast to the case of the molecular dynamics simulations, where the unitary directly represented the physical temporal evolution of the molecule and the photonic state the initial vibrational configuration of the molecule.

- No post-selection. In all boson sampling protocols, and also in the simulation of molecular dynamics, the detection was conditional on events with a fixed number of photons, n. In the calculation of the Franck-Condon profiles, as in eq. (5.2.9), all possibilities $n \in [0, \infty)$ are present, which means that no post-selection has to be performed and all events have to be stored (including those where no photons were detected). Moreover, terms where multiple photons emerge from the same mode are in general relevant, meaning that number-resolving detection has to be adopted.

- Reconstruction of the probability distribution. In contrast to sampling protocols, where we are concerned in sampling from the output distribution, we are here interested in the reconstructing the full Franck-Condon profile, i.e. the full output distribution $p_{\mathrm{FC}}(\boldsymbol{k})$. However, at each run we collapse the output state via the photon detection, providing a single sample from the distribution. We will have to run the protocol multiple times to reconstruct the total distribution up to a statistical error small enough. However, the central limit theorem implies that the average number of samples required to estimate a Franck-Condon coefficient $\mathrm{FCP}(\omega)$ up to an error ϵ scales as $\mathcal{O}(\epsilon^{-2})$, which only represents a polynomial overhead in the protocol run-time [3].

While the formalism of Gaussian boson sampling could be generalised and used to calculate output probabilities $p_{\mathrm{FC}}(\boldsymbol{k}) = |\langle \boldsymbol{k}| \mathcal{U}_{\mathrm{L}} |\psi\rangle|^2$ also for Gaussian input states with non-zero displacement [24, 25], for convenience we will focus the discussion on cases with no displacement[5]. In this situation we have $\alpha_i = 0$ for all inputs, and we can drop the unitary operation $\mathcal{U}_{\mathrm{R}}$ in eq. (5.2.12) as it acts on the vacuum. The input state is then now a set of single-mode squeezed states $|\psi\rangle = \oplus_{j=1}^{m} |\mathrm{SMS}(\xi_i)\rangle_j$ (with $|\mathrm{SMS}(\boldsymbol{\xi})\rangle$ defined as in eq. 4.1.18), and the experimental configuration, shown in Fig. 5.5b is very similar to the Gaussian boson sampling scheme discussed in chapter 4. In particular, the Franck-Condon factors can now be explicitly expressed in terms of the Hafnian matrix function as

$$p_{\mathrm{FC}}(\mathbf{k}) = \frac{|\mathrm{Haf}(B_{\mathbf{k}})|^2}{k_1! \cdots k_m! \sqrt{\det(\sigma_{\mathrm{Q}})}}, \tag{5.2.13}$$

where $B_{\mathbf{k}}$ is obtained from $B = U_{\mathrm{L}} \cdot \oplus_j (\tanh \xi_j) \cdot U_{\mathrm{L}}^{\top}$ by repeating k_i times its ith column and row, and σ_{Q} is given in eq. (4.1.14).

5 Note that molecular structural changes which do not need displacement do exist. An example is the the tropolone molecule, $C_7H_6O_2$, whose FC calculation requires only squeezed vacuum states [26]. A zero-displacement description can thus be sufficient in particular cases.

5.2.3 *Experimental test on a silicon device*

The close similarity between the algorithm to reconstruct Franck-Condon profiles and Gaussian boson sampling allowed us to test it on the silicon photonic device reported in chapter 4. We considered a simplified zero-displacement scenario, with input squeezed states generated via degenerate spontaneous four-wave mixing, using the dual-pump scheme shown in Fig. 4.8b.

In contrast to our implementation of Gaussian boson sampling, to implement the photonic Franck-Condon profile calculation we needed to resolve the photon number at the photon counter. This was performed by adopting a pseudo-number-resolving detection scheme where up to 2 photons could be probabilistically resolved in the output modes. This can be achieved with a slight modification at the detection apparatus: we inserted 50:50 fibre beam splitters on the output fibres and detected all the single outputs with the superconducting single photon detectors, as pictured in Fig. 5.6. In this way, if two photons are present in a single output mode, they are probabilistically split at the beam splitter with a probability one-half, and the simultaneous detection of both the associated detectors identifies the presence of two photons in the mode. The cases where a single photon is in a mode are not affected, as it is detected by either one of the two detectors associated to this mode.

The operation U_{L} is performed in the evolution through the random walk (see Fig. 4.8a). As the quantum walk embedded in the chip cannot be reconfigured, we can't tune the implemented unitary to match a particular target molecule. Nevertheless, we can think of reversing the process. In particular, we can consider a synthetic molecule whose vibronic transition matched the implemented unitary, the FC profile calculation protocol as a benchmarking tool to test the performance of near-term devices in this application. The merit of such a procedure is to understand the fidelities that one could expect to find in a similar setting engineered to investigate a particular molecule in the silicon quantum photonics technology.

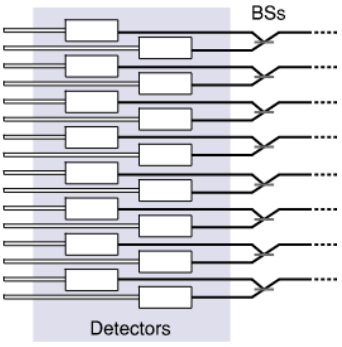


Figure 5.6: Detection schematic for pseudo-photon number resolving measurements. Pseudo number resolving is obtained by probabilistically splitting the incoming photons via on-fibre 1×2 beam-splitters and using SNSPDs at the outputs.

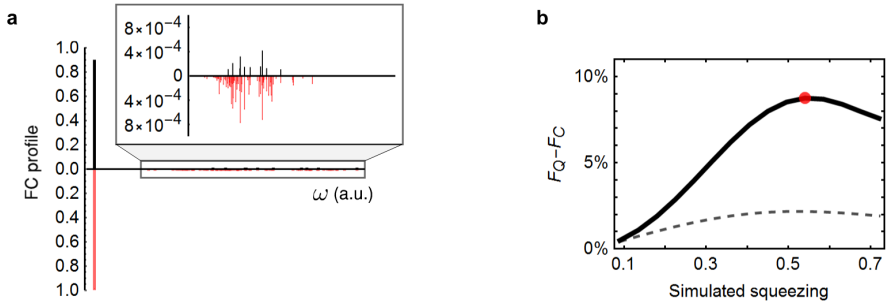


Figure 5.7: **a**, Reconstructed Franck-Condon factors for a synthetic molecule directly associated to the device parameters. **b**, Improvement over optimal classical strategies in the fidelity of estimating the Franck-Condon profile, using post-processing to simulate higher squeezing. The solid line plots cases that simulate squeezing by processing experimental data, while the dashed line indicates no post-processing.

To quantify the advantage of the quantum experiment over classical approximation strategies we adopted the approach developed by Clements et al. [26]. The idea is to quantify this advantage in terms of the fidelity of experimentally reconstructed profiles with respect to the fidelity obtainable by the best classical approximation. The latter represents the Franck-Condon profile with highest fidelity achievable from classical experiments employing classical input states only[6]. As the fidelity between states is invariant under unitary evolutions, finding the optimal classical strategy corresponds to finding the closest classical state to the initial displaced squeezed state [26]. In turn, the single-mode classical state closest to a single-mode displaced Gaussian state is a coherent state with the same level of displacement [28]. Consequently, in absence of displacement, vacuum is the optimal classical state. To quantify the quantum advantage obtained in the experiment we thus consider the difference $\mathcal{C} = F_{\mathrm{Q}} - F_{\mathrm{C}}$ between the fidelity F_{Q} of the Franck-Condon profile reconstructed from the experiment to the theoretical one, and the fidelity obtained using the optimal classical strategy F_{C}.

As a first test, we reconstructed the Franck-Condon profile for the synthetic molecule directly associated to the device, i.e. with a vibronic transition described by the squeezing parameters $\boldsymbol{\xi}$ and the unitary U_{L} implemented in our device. As the choice of frequencies $\{\omega_i'\}$ does not affect the fidelity of an Franck-Condon profile (only its shape), which is the parameter of interest here, we choose them randomly and with arbitrary unit of measure. This defines the synthetic molecule used for the benchmarking. In this case, although the fidelity of the profile is very high ($F_{\mathrm{Q}} > 99\%$), the quantum advantage is only $\mathcal{C} = 0.4\%$, i.e. while there is still a quantum advantage, it is actually very low. This is due to the

6 More in detail, the definition of classical input states used in Ref. [26] is states with non-negative Glauber-Sudarshan function [27].

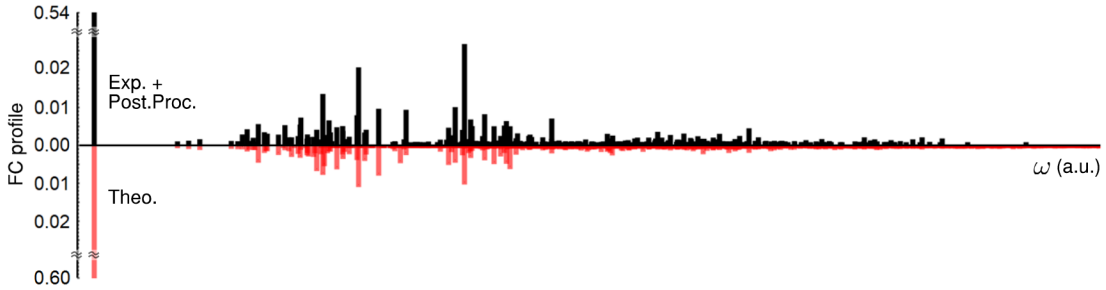


Figure 5.8: Reconstructed Franck-Condon factors calculated from processing measured data (black) for frequencies ω (in arbitrary units). The plot is for a simulated squeezing corresponding to $\gamma = 5$ (red point in Fig. 5.7b). Theoretical estimates for contributions of up to 8 photons (> 8 photon contributions are negligible) are reported in black lines.

fact that the squeezing in the sources is very low ($\lesssim 0.2$), which means that, as shown in Fig. 5.7a the Franck-Condon profile is dominated by classical vacuum contributions.

To obtain higher quantum advantages, molecules involving higher squeezing values have to be involved. This can be simulated to some extent via data post-processing, based on the characterised losses and detection efficiency (see section 4.2.3), using the following prescription. For simplicity, we focus on the case where no displacement is used, although the approach can be easily generalised to the case with non-zero displacement. Consider the situation where we want to reconstruct the Franck-Condon profile $p_{\rm FC}(\boldsymbol{k})$ for a molecule with corresponding squeezing parameters $\boldsymbol{\xi}$ and unitary operation $U_{\rm L}$. However, in our experiment we have only access to a limited amount of squeezing, $\bar{\xi}_{\max} \leq \max_i \xi_i$, so that for each source we can only implement a squeezing $\bar{\xi}_i \leq \bar{\xi}_{\max}$. The transition amplitudes can be then estimated in the imperfect device by tuning the squeezing parameters in the sources so that $\tanh(\bar{\xi}_i) = \gamma \tanh(\xi_i)$, with $0 < \gamma \leq 1$ a real constant rescaling, so that $[\oplus_i \tanh(\bar{\xi}_i)] = \gamma[\oplus_i \tanh(\xi_i)]$. In this way, considering also uniform losses η in the device, from eq.(5.2.13), we have a simple relation between the ideal Franck-Condon factors $p_{\rm FC}(\boldsymbol{k})$ and the ones reconstructed in the imperfect device $\overline{p}_{\rm FC}(\boldsymbol{k})$, given by

$$\overline{p}_{\rm FC}(\boldsymbol{k}) = \mathcal{N}\eta^n \gamma^{2n} p_{\rm FC}(\boldsymbol{k}), \tag{5.2.14}$$

with $n = \sum_{i=1}^m k_i$ the total number of photons in the output configuration $\boldsymbol{k}$, and $\mathcal{N}$ a normalisation constant. As η and γ are known from the characterisation of the device, eq. (5.2.14) can be inverted to reconstruct the Franck-Condon profile via

$$p_{\rm FC}(\boldsymbol{k}) = \frac{\overline{p}_{\rm FC}(\boldsymbol{k})/\eta^{\sum k_i}\gamma^{2\sum k_i}}{\sum_{\boldsymbol{x}} \overline{p}_{\rm FC}(\boldsymbol{x})/\eta^{\sum x_i}\gamma^{2\sum x_i}}. \tag{5.2.15}$$

Note that the computational complexity of the experiment is not increased by doing this data post-processing, in the sense that, if from the experiment we are able to collect only up to n-photon events, then we are truncating the Franck-Condon profile in eq. (5.2.9) at n-photon contributions. This implies that a Franck-Condon profile reconstructed with this approach is faithful only if the ideal contributions involving more than n photons are negligible, which, for large values of n, is in general true only for η and γ close to unity. Therefore, the approach is clearly not scalable, i.e. it will not provide much help in future large-scale implementations. Nonetheless, it can provide significant advantages in near-term quantum devices.

We implemented the post-processing procedure described above to reconstruct Franck-Condon profiles of synthetic molecules requiring higher squeezing. The synthetic molecule to be simulated is obtained as in the first test above, with the only difference that now the associated squeezing values are $\xi_i = \bar{\xi}_i/\gamma$, where $\bar{\xi}_i$ are the squeezing parameters in the device, preliminarily characterised (see section 4.2.3). In Fig. 5.7b the quantum advantage over optimal classical strategies is reported for different values of simulated squeezing (reported as $\max_i \xi_i$), that is for different values of γ, both for the case where the post-processing is performed (solid black line) and when raw data is instead used (dashed black line). It can be noted that the procedure provides a significant improvement, with the quantum advantage initially increasing and then decreasing when the Franck-Condon terms, due to more than 4-photon contributions (the highest number of photons detected in the experiment), start to be dominant. The maximum quantum advantage $\mathcal{C} = 9\%$ is obtained for $\gamma \approx 5$, for which the Franck-Condon profile, plotted in Fig. 5.8, has a fidelity of 86% with the ideal one. In all cases, applying the post-processing procedure provides a significant advantage as compared to raw data usage.

5.3 Discussion

Mappings between both vibrational dynamics and vibronic transitions in molecules with photonic boson sampling protocols have been described. The latter has been experimentally tested on a silicon photonic circuit with the on-chip generation and processing of weak squeezed-vacuum quantum states and pseudo-number-resolving detection.

Supported by the technological developments described in previous chapters, these quantum simulation protocols are promising for achieving quantum speed-ups in industrially relevant applications. For example, the ability to have reconfigurable photonic devices acting as reprogrammable molecules could be applied to design reaction processes with lower energy consumption [2]. It is however important to identify what are the problems of importance in chemistry that truly require the use of quantum simulators. Close and interdisciplinary interaction with quantum chemists will be key to establish the

perspectives of the approach. For instance, for the calculation of Franck-Condon spectra, it is not clear whether situations of interest will need configurations with more than 20 photons. It could be that classical approaches to simulate boson sampling machines could be enough to address all instances of practical relevance.

Another issue is that, for molecules with a high degree of complexity, the harmonic approximation can potentially cease to be accurate. To maintain a good fidelity, inharmonicities will need to be included in the simulation. In Ref. [2] it was shown that, for the vibrational dynamics simulation, inharmonicities can be inserted in the mapping by including optical non-linearities in the optical circuit. This can potentially be quite experimentally demanding.

Finally, another interesting difference between pure boson sampling and the quantum chemistry applications based on it is that quantum simulation tasks allow a straightforward verification approach. Namely, build the simulated molecule to see if the statistics measured on the quantum hardware are consistent with the actual physical system.

Acknowledgements

The theoretical frameworks for mapping of vibrational dynamics and vibronic transitions were originally developed in Ref. [2] (together with experimental demonstration) and Ref. [3], respectively. The author's contributions to the material reported in this chapter are the implementation of the algorithm for the Franck-Condon profile reconstruction using the silicon device reported in chapter 4, the data analysis, and the development of the post-processing procedure to simulate Franck-Condon profiles for higher values of squeezing. The work is published in Ref. [1]. Support was provided by Dr. Levon Chakhmakhchyan and Dr. Anthony Laing for the data analysis. We are also grateful to them and to Ms. Rachel Chadwick, Mr. Nicola Maraviglia, and Dr. Raffaele Santagati for useful discussions on the mappings.

References

1. Paesani, S., Ding, Y., Santagati, R., Chakhmakhchyan, L., Vigliar, C., Rottwitt, K., Oxenløwe, L. K., Wang, J., Thompson, M. G. & Laing, A. Generation and sampling of quantum states of light in a silicon chip. *Nature Physics* **15,** 925–929 (2019).
2. Sparrow, C., Martín-López, E., Maraviglia, N., Neville, A., Harrold, C., Carolan, J., Joglekar, Y. N., Hashimoto, T., Matsuda, N., O'Brien, J. L., *et al.* Simulating the vibrational quantum dynamics of molecules using photonics. *Nature* **557,** 660 (2018).
3. Huh, J., Guerreschi, G. G., Peropadre, B., McClean, J. R. & Aspuru-Guzik, A. Boson sampling for molecular vibronic spectra. *Nature Photonics* **9,** 615 (2015).
4. González Olivares, D., Peropadre, B., Aspuru-Guzik, A. & García-Ripoll, J. J. Quantum simulation with a boson sampling circuit. *Phys. Rev. A* **94,** 022319. https://link.aps.org/doi/10.1103/PhysRevA.94.022319 (2016).
5. Arrazola, J. M. & Bromley, T. R. Using Gaussian Boson Sampling to Find Dense Subgraphs. *Phys. Rev. Lett.* **121,** 030503. https://link.aps.org/doi/10.1103/PhysRevLett.121.030503 (2018).
6. Brádler, K., Dallaire-Demers, P.-L., Rebentrost, P., Su, D. & Weedbrook, C. Gaussian Boson Sampling for perfect matchings of arbitrary graphs. *arXiv:1712.06729* (2017).
7. Brádler, K., Friedland, S., Izaaca, J., Killoran, N. & Su, D. Graph isomorphism and Gaussian boson sampling. *arXiv:1810.10644* (2018).
8. Dirac, P. A. M. *The principles of quantum mechanics* (Oxford University Press, Oxford, 1981).
9. Feynman, R. Simulating Physics with Computers. *Int. J. Theor. Phys.* **21,** 467 (1982).
10. Born, M. & Oppenheimer, R. Zur quantentheorie der molekeln. *Annalen der Physik* **389,** 457–484 (1927).

11. Lewars, E. in *Computational Chemistry: Introduction to the Theory and Applications of Molecular and Quantum Mechanics* (2nd edn. Springer, New York, 2011).

12. Friese, D. H., Hättig, C. & Koßmann, J. Analytic Molecular Hessian Calculations for CC2 and MP2 Combined with the Resolution of Identity Approximation. *Journal of Chemical Theory and Computation* **9,** 1469–1480. https://doi.org/10.1021/ct400034t (2013).

13. Hundt, P. M., Jiang, B., van Reijzen, M. E., Guo, H. & Beck, R. D. Vibrationally promoted dissociation of water on Ni (111). *Science* **344,** 504–507 (2014).

14. Carolan, J., Harrold, C., Sparrow, C., Martín-López, E., Russell, N. J., Silverstone, J. W., Shadbolt, P. J., Matsuda, N., Oguma, M., Itoh, M., Marshall, G. D., Thompson, M. G., Matthews, J. C. F., Hashimoto, T., O'Brien, J. L. & Laing, A. Universal linear optics. *Science* **349,** 711–716 (2015).

15. Hachmann, J., Olivares-Amaya, R., Atahan-Evrenk, S., Amador-Bedolla, C., Sánchez-Carrera, R. S., Gold-Parker, A., Vogt, L., Brockway, A. M. & Aspuru-Guzik, A. The Harvard clean energy project: large-scale computational screening and design of organic photovoltaics on the world community grid. *The Journal of Physical Chemistry Letters* **2,** 2241–2251 (2011).

16. Gross, M., Müller, D. C., Nothofer, H.-G., Scherf, U., Neher, D., Bräuchle, C. & Meerholz, K. Improving the performance of doped π-conjugated polymers for use in organic light-emitting diodes. *Nature* **405,** 661 (2000).

17. Santoro, F., Lami, A., Improta, R. & Barone, V. Effective method to compute vibrationally resolved optical spectra of large molecules at finite temperature in the gas phase and in solution. *The Journal of chemical physics* **126,** 184102 (2007).

18. Jankowiak, H.-C., Stuber, J. & Berger, R. Vibronic transitions in large molecular systems: Rigorous prescreening conditions for Franck-Condon factors. *The Journal of chemical physics* **127,** 234101 (2007).

19. Hayes, D., Wen, J., Panitchayangkoon, G., Blankenship, R. E. & Engel, G. S. Robustness of electronic coherence in the Fenna–Matthews–Olson complex to vibronic and structural modifications. *Faraday discussions* **150,** 459–469 (2011).

20. Duschinsky, F. The importance of the electron spectrum in multi atomic molecules. Concerning the Franck- Condon principle. *Acta Physicochim. URSS* **7,** 551 (1937).

21. Franck, J & Dymond, E. Elementary processes of photochemical reactions. *Trans. Faraday Soc.* **21,** 536 (1926).

22. Condon, E. Nuclear motions associated with electron transitions in diatomic molecules. *Phys. Rev.* **32,** 858 (1928).

23. Doktorov, E., Malkin, I. & Man'ko, V. I. Dynamical symmetry of vibronic transitions in polyatomic molecules and the Franck-Condon principle. *J. Mol. Spectrosc.* **64,** 302 (1977).

24. Hamilton, C. S., Kruse, R., Sansoni, L., Barkhofen, S., Silberhorn, C. & Jex, I. Gaussian Boson Sampling. *Phys. Rev. Lett.* **119,** 170501. https://link.aps.org/doi/10.1103/PhysRevLett.119.170501 (2017).

25. Kruse, R., Hamilton, C. S., Sansoni, L., Barkhofen, S., Silberhorn, C. & Jex, I. A detailed study of Gaussian Boson Sampling. *arXiv:1801.07488* (2018).

26. Clements, W. R., Renema, J. J., Eckstein, A., Valido, A. A., Lita, A., Gerrits, T., Nam, S. W., Kolthammer, W. S., Huh, J. & Walmsley, I. A. Experimental Quantum Optical Approximation of Vibronic Spectroscopy. *arXiv:1710.08655* (2017).

27. Leonhardt, U. *Essential Quantum Optics* (Cambridge University Press, 2010).

28. Marian, P., Marian, T. A. & Scutaru, H. Quantifying Nonclassicality of One-Mode Gaussian States of the Radiation Field. *Phys. Rev. Lett.* **88,** 153601. https://link.aps.org/doi/10.1103/PhysRevLett.88.153601 (2002).

6 DIGITAL QUANTUM SIMULATION WITH PHOTONS

Let us now focus on a different approach to quantum simulation, based on digital quantum information processing. While in analog quantum simulation the simulator directly emulates a specific real system of interest, digital quantum simulation maps the target problem onto a set of gates which can be implemented by a quantum computer. An important difference between these two types of quantum simulation is that, while the underlying theory for analog methods strongly depends on the platform on which the simulated system is mapped into (e.g. photonics, ultra-cold atoms, QED systems, etc.), digital algorithms are completely platform-independent. Digital protocols essentially consist of a list of gates that need to be applied to qubits, and are not concerned what the qubits are made of, or what type of interaction is used to perform the operations. As a consequence, although "with photons" is included in the title of this chapter, most of the algorithms that will be discussed are general and do not assume a-priori a photonic implementation. We will, however, show how photonic circuits can be designed to embed simple digital quantum circuits in reconfigurable integrated photonic chips to implement the digital quantum simulation protocols discussed in this chapter.

Due to their platform-independence, digital approaches are more general than analog methods. On the other hand, the hardware requirements for implementing digital protocols at a scale that can challenge classical approaches are much more demanding compared to analog ones. In particular, error correction of the digital circuits is believed to be required in such instances, with overheads that make the implementation of large-scale digital quantum simulators likely to be far beyond current experimental capabilities in any platform. The development of digital algorithms with increased robustness to noise is thus crucial to decrease such overheads as much as possible, and currently represents an important field of research.

We will focus on a description of two different types of algorithms for digital quantum simulation: simulations based on quantum phase estimation, and quantum variational eigensolvers. For the first case, we will report a Bayesian technique which enables noise-robustness, a notorious challenge for quantum phase estimation methods. For the second case, an approach for the simulation of excited states will be described. For both types of quantum algorithms, a proof-of-principle demonstration on a silicon photonic chip will be reported.

6.1 Electrons, molecules, and qubits

In chapter 5 we described mappings between quantum chemistry problems concerning the nuclear vibrations of molecules into analog photonic systems. Using the Born-Oppenheimer approximation, we supposed the vibrational dynamics as decoupled from the electronic degrees of freedom, and considered the electronic structure and the potential energy surface as inputs into the problem. Most of the algorithms developed for quantum chemistry simulation on a quantum computer are instead concerned with the problem of finding the electronic structure itself.

6.1.1 *Electronic structures in quantum chemistry*

The Born-Oppenheimer approximation essentially states that, as electronic masses are orders of magnitude smaller than those of the nuclei, the dynamics of the electrons and nuclei happen at different timescales [3]. This means that we can consider the nuclei to be stationary and thus have time-independent electronic dynamics. Mathematically, the approximation consists in considering a total wavefunction $\Psi = \psi_e \times \psi_n$ where the electronic (ψ_e) and nuclear (ψ_n) contributions are uncorrelated. The problem of finding the electronic structure is then reduced to solving the time-independent Schrödinger equation

$$i\hbar\frac{\partial}{\partial t}\psi_e = \hat{H}_e\psi_e. \tag{6.1.1}$$

The explicit form of the electronic Hamiltonian $\hat{H}_e$, for a system of N electrons and M nuclei, is given by

$$\hat{H}_e = -\sum_i \frac{\nabla_i^2}{2} - \sum_{i,I} \frac{Z_I}{|\boldsymbol{r}_i - \boldsymbol{R}_I|} + \sum_{i\neq j} \frac{1}{2\,|\boldsymbol{r}_i - \boldsymbol{r}_j|}, \tag{6.1.2}$$

where $\boldsymbol{r}_i$ is the position of the i-th electron ($i \in \{1, \ldots, N\}$), and $\boldsymbol{R}_I$ and Z_I are the position and charge of the I-th nucleus ($I \in \{1, \ldots, M\}$). The first term represents the

kinetic contributions of the electrons, while the second and third terms are the Coulomb interactions between nuclei and electrons, and between electrons, respectively.

In order to have a second-quantization description of the electronic Hamiltonian $\hat{H}_e$ it is convenient to introduce a basis for the electronic wavefunctions

$$\{\phi_p(\boldsymbol{r}_i)\}_{p\in\{1,\dots,M\}} \tag{6.1.3}$$

given by approximate electron molecular spin-orbitals. Numerous quantum chemistry methods can be used to calculate approximate orbitals $\phi_p(\boldsymbol{r}_i)$ through different approaches (see e.g. Ref. [4]). We can then define the electron creation $\hat{a}_p^\dagger$ and annihilation $\hat{a}_p$ operators associated to the orbital ϕ_p, which are subject to the fermionic anticommutation relations

$$\begin{aligned} \{\hat{a}_p, \hat{a}_q^\dagger\} &= \delta_{pq}\hat{\mathbb{1}}, \\ \{\hat{a}_p, \hat{a}_q\} &= 0, \end{aligned} \tag{6.1.4}$$

where $\{A, B\} = AB + BA$ is the anticommutator. In particular, these rules imply $(\hat{a}_q^\dagger)^2 = (\hat{a}_q)^2 = 0$. Once we have defined these operators we can write the Hamiltonian H_e in eq.(6.1.2) via projecting it into the basis $\{\phi_p\}$, obtaining

$$\hat{H}_e = \sum_{pq} h_{pq}\hat{a}_p^\dagger\hat{a}_q + \sum_{pqrs} h_{pqrs}\hat{a}_p^\dagger\hat{a}_q^\dagger\hat{a}_r\hat{a}_s. \tag{6.1.5}$$

The values of h_{pq} and h_{pqrs}, which are explicitly given by

$$h_{pq} = \int \mathrm{d}\boldsymbol{r}\ \phi_p^*(\boldsymbol{r})\left(-\frac{\nabla_i^2}{2} - \sum_I \frac{Z_I}{|\boldsymbol{r}_i - \boldsymbol{R}_I|}\right)\phi_q(\boldsymbol{r}) \tag{6.1.6}$$

$$h_{pqrs} = \int \mathrm{d}\boldsymbol{r}_1\mathrm{d}\boldsymbol{r}_2\ \frac{\phi_p^*(\boldsymbol{r}_1)\phi_q^*(\boldsymbol{r}_2)\phi_s(\boldsymbol{r}_1)\phi_r(\boldsymbol{r}_2)}{|\boldsymbol{r}_i - \boldsymbol{r}_j|}, \tag{6.1.7}$$

depend on the nuclear structure (which is time-independent in the Born-Oppenheimer approximation) of the particular molecule under study, and, from here on, we will consider them as inputs to the problem. The electronic structure problem then consists in calculating the energy $E = \langle\lambda_E|\,\hat{H}_e\,|\lambda_E\rangle$ corresponding to a chemically relevant electronic eigenstate $|\lambda_E\rangle$ of the Hamiltonian $\hat{H}_e$, e.g. the ground-state or some low-energy excited state. For example, the potential energy landscapes used to study molecular vibrational dynamics in chapter 5 can correspond to the ground-energy of the target molecule as a function of the nuclei coordinates $\boldsymbol{R}_I$.

6.1.2 *Description of electronic wavefunctions*

A first issue we encounter when tackling the electronic structure problem is to find an efficient way to describe the states $|\lambda_E\rangle$. The first property that has to be taken into account is the Pauli exclusion principle, which requires the electronic wavefunction to be anti-symmetric:

$$\psi(\ldots,\boldsymbol{r}_i,\ldots,\boldsymbol{r}_j,\ldots) = -\psi(\ldots,\boldsymbol{r}_j,\ldots,\boldsymbol{r}_i,\ldots) \tag{6.1.8}$$

for arbitrary $i \neq j$. For configurations with N electrons and M nuclei, where orbitals $\{p_1,\ldots,p_N\}$ are occupied, states that satisfy this condition can be conveniently written in terms of the single-electron basis $\{\phi_p\}$ via the Slater determinant:

$$\psi(\boldsymbol{r}_1,\boldsymbol{r}_1,\ldots,\boldsymbol{r}_N,) = \frac{1}{\sqrt{N!}}\begin{vmatrix} \phi_{p_1}(\boldsymbol{r}_1) & \phi_{p_2}(\boldsymbol{r}_1) & \cdots & \phi_{p_N}(\boldsymbol{r}_1) \\ \phi_{p_1}(\boldsymbol{r}_2) & \phi_{p_2}(\boldsymbol{r}_2) & \cdots & \phi_{p_N}(\boldsymbol{r}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{p_1}(\boldsymbol{r}_N) & \phi_{p_N}(\boldsymbol{r}_N) & \cdots & \phi_{p_N}(\boldsymbol{r}_N) \end{vmatrix}. \tag{6.1.9}$$

In the second-quantisation formalism, the Slater determinants can be simply represented in terms of fermionic Fock states

$$\psi(\boldsymbol{r}_1,\boldsymbol{r}_1,\ldots,\boldsymbol{r}_N,) \mapsto |\mathbf{n}\rangle = |n_1,n_2,\ldots,n_M\rangle = \prod_{p=1}^{M}(\hat{a}_p^\dagger)^{n_p}\,|\mathrm{vac}\rangle\,, \tag{6.1.10}$$

where $n_p \in \{0,1\}$ is the electronic occupation number of the orbital ϕ_p (with $\sum_p n_p = N$), and $\hat{a}_p^\dagger$ is the fermionic creation operator associated to ϕ_p, as defined above. The action of the creation and annihilation operators on the state $|\mathbf{n}\rangle$ are given by:

$$\begin{aligned} \hat{a}_p^\dagger\,|\mathbf{n}\rangle &= \delta_{n_p,0}(-1)^{\sum_{i=1}^{p-1} n_i}\left|n_1,n_2,\ldots,n_p'=1,\ldots,n_M\right\rangle, \\ \hat{a}_p\,|\mathbf{n}\rangle &= \delta_{n_p,1}(-1)^{\sum_{i=1}^{p-1} n_i}\left|n_1,n_2,\ldots,n_p'=0,\ldots,n_M\right\rangle. \end{aligned} \tag{6.1.11}$$

In second-quantization, transformations between different Slater determinants can then be easily achieved by applying combinations of creation and annihilation operators. For example, starting from an initial reference Slater determinant $|\bar{\mathbf{n}}\rangle$, all determinants that differ from $|\bar{\mathbf{n}}\rangle$ by a single occupied orbital ($\bar{p}_i \to u_i$, same total number of electrons) can be written as $\hat{a}_{u_i}^\dagger \hat{a}_{\bar{p}_i}\,|\bar{\mathbf{n}}\rangle$. Those differing by two orbitals are of the form $\hat{a}_{u_i}^\dagger \hat{a}_{v_j}^\dagger \hat{a}_{\bar{p}_i}\hat{a}_{\bar{q}_j}\,|\bar{\mathbf{n}}\rangle$, and so on. Proceeding in the same manner we can construct the full set of Slater determinants $\{|\mathbf{n}\rangle\}$, which forms a complete basis of the Hilbert space spanned by

systems with N electrons and M nuclei [4]. Arbitrary eigenstates of the Hamiltonian $\hat{H}_e$ can then be written as a linear combination of Slater determinants

$$|\psi\rangle = \sum_{\mathbf{n}} c_{\mathbf{n}} |\mathbf{n}\rangle \tag{6.1.12}$$

$$= \left(\hat{\mathbb{1}} + \sum_{u,\bar{p}} c_{u,\bar{p}} \hat{a}_u^\dagger \hat{a}_{\bar{p}} + \sum_{u,v,\bar{p},\bar{q}} c_{u,v,\bar{p},\bar{q}} \hat{a}_u^\dagger \hat{a}_v^\dagger \hat{a}_{\bar{p}} \hat{a}_{\bar{q}} + \dots \right) |\bar{\mathbf{n}}\rangle \,, \tag{6.1.13}$$

where the indices $u, v, \dots$ run over all orbitals unoccupied in $|\bar{\mathbf{n}}\rangle$, and the indices $\bar{p}, \bar{q}, \dots$ over all occupied orbitals. This formulation, known as *full configuration interaction* (FCI) provides an exact description of the wavefunction. However, as the number of ways to occupy N of the M orbitals is $\binom{M}{N}$, the number of parameters that enter in eq. (6.1.12) and in eq. (6.1.13) increases exponentially with N. This overhead makes the FCI representation impractical for complex molecules. For example, when trying to find an eigenstate of interest, we would need to search through an exponentially large parameter space, which becomes rapidly intractable.

In quantum chemistry one route to avoid this issue and make the wavefunction description tractable is to use approximations which consider a reduced, polynomially large, number of Slater determinants, known as *ansätze*.

Hartree-Fock

The most simple approach, which is widely used in classical techniques, is the Hartree-Fock method. As we will see, it is an approximation of the electronic wavefunction which uses a single Slater determinant. More in detail, it represents a mean-field approximation of the electronic wavefunction, where the underlying idea is to start from the set of N wavefunctions $\{\phi^{(0)}(\mathbf{r}_i)\}$ of N independent electrons obtained by neglecting the electron-electron interaction term in eq.(6.1.2). We then proceed iteratively, where at the k-th step we assume each electron to move in the average charge distribution of all of the other electrons according to $\{\phi^{(k-1)}(\mathbf{r}_i)\}$, which introduces an effective potential. The new set $\{\phi^{(k)}(\mathbf{r}_i)\}$ is calculated solving again a single-electron eq.(6.1.2) where the effective potential is used in place of electron-electron interaction term. This procedure is repeated until the orbitals converge to the set of N orbitals $\{\phi(\mathbf{r})\}$. The Hartree-Fock wavefunction $\psi_{\text{HF}}(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) \mapsto |\psi_{\text{HF}}\rangle$ is finally obtained by taking the Slater determinant of the resulting molecular orbitals. The resulting wavefunction $|\psi_{\text{HF}}\rangle$ can be shown to be the wavefunction consisting of a single Slater determinant with the lowest energy [5], and can thus be used as an approximation of the ground-state.

The Hartree-Fock method has been very successful in describing a broad range of molecular systems, and can be efficiently performed on classical computers [5]. However, consisting of a mean-field approximation, it provides only a limited description of the electron-electron correlations. As a result, the Hartree-Fock approach becomes inaccurate when applied to strongly correlated molecular systems [5]. This limitation has induced the development of so called post-Hartree-Fock methods, where an initial reference state, typically the Hartree-Fock state, is further processed to re-introduce electron correlations. Some examples of post-Hartree-Fock methods are described below.

Coupled cluster

So far, we have first discussed the full configuration interaction approach, which provided an exact description but included exponentially many Slater determinants, making calculations intractable. In contrast, the Hartree-Fock approach considers only a single determinant, and is thus computationally accessible, but provides an unreliable approximation in the presence of strong correlations. The idea is now to explore the regime in between these two cases: descriptions which involve multiple, but only polynomially many, Slater determinants.

A first approach, called *configuration interaction*, uses the Hartree-Fock ansatz as reference state in eq. (6.1.13), and truncates the series to a small number of excitations. The reasoning behind this is that, when targeting the ground-state or low-energy excited states, low-energy excitations dominate the electronic wavefunction, which can thus be well approximated by just considering the first terms in the linear parametrisation in eq. (6.1.13). However, using this type of parametrisation, the convergence to the full configuration interaction is slow and impractical for strongly correlated systems [5].

The state-of-the-art for classically tractable ansätze is the *coupled cluster* method. Closely related to the configuration interaction, it uses a parametrisation which can be expressed in an exponential form:

$$|\psi_{\mathrm{CC}}\rangle = e^{\hat{T}(\boldsymbol{\theta})} |\psi_{\mathrm{HF}}\rangle \,, \tag{6.1.14}$$

where

$$\hat{T}(\boldsymbol{\theta}) = \sum_i \hat{T}_i(\boldsymbol{\theta}) \tag{6.1.15}$$

$$\hat{T}_1(\boldsymbol{\theta}) = \sum_{u,\bar{p}} \theta_{u,\bar{p}} \hat{a}_u^\dagger \hat{a}_{\bar{p}}, \tag{6.1.16}$$

$$\hat{T}_2(\boldsymbol{\theta}) = \sum_{u,v,\bar{p},\bar{q}} \theta_{u,v,\bar{p},\bar{q}} \hat{a}_u^\dagger \hat{a}_v^\dagger \hat{a}_{\bar{p}} \hat{a}_{\bar{q}}, \tag{6.1.17}$$

$$\ldots \text{and so on.} \tag{6.1.18}$$

Including all excitation operators $\hat{T}_i$ we recover the full configuration interaction scenario. Because of its exponential parametrisation, even when the operator series at the exponent is truncated at $i \leq i_{\max}$ for some $i_{\max}$ (usually single and double excitations are considered $\hat{T} \approx \hat{T}_1 + \hat{T}_2$), the coupled cluster ansatz still generates a trial wavefunction which includes all possible Slater determinants, albeit with an imperfect parametrisation respect to the full configuration interaction. This property allows us to obtain better approximations with respect to the configuration interaction approach and needs fewer parameters. However, it still struggles to describe strong correlations, and the non-unitary form of the operator $\exp(\hat{T})$ can make it impractical to implement classically [4–6].

Unitary coupled cluster

The limitations of the coupled cluster approach can be overcome by redefining the exponential operator in a unitary form, and formulating the ansatz as:

$$|\psi_{\text{UCC}}\rangle = e^{\hat{T}(\boldsymbol{\theta}) - \hat{T}^\dagger(\boldsymbol{\theta})} |\psi_{\text{HF}}\rangle \tag{6.1.19}$$

$$= U_{\text{UCC}}(\boldsymbol{\theta}) |\psi_{\text{HF}}\rangle \,, \tag{6.1.20}$$

with $\hat{T}(\boldsymbol{\theta})$ defined as for the coupled cluster approach in eq. (6.1.15). This method is called *unitary coupled cluster*. In contrast to the standard couple cluster approach, the operator $\hat{T} - \hat{T}^\dagger$ at the exponent is now manifestly anti-Hermitian, implying that the exponential unitary coupled cluster operator $U_{\text{UCC}} = \exp(\hat{T} - \hat{T}^\dagger)$ is unitary. As a consequence, it can be mapped into a true evolution of a quantum system initialised in the state $|\psi_{\text{HF}}\rangle$, and can be then treated in a variational way [6]. The ansatz is typically truncated at single or double excitation level, which allows us to describe the wavefunction via a polynomial ($\mathcal{O}(M^2N^2)$ number of parameters (see [4]).

While the unitary coupled cluster approach maintains all of the advantages of the coupled cluster method with additional benefits, e.g. practical for variational approaches, the necessity to operate U_{UCC} on exponentially large (in N) state vectors makes it still intractable on classical computers [4, 6]. However, as we will see in section 6.2.3, the operation $U_{\text{UCC}}(\boldsymbol{\theta})$ can be efficiently decomposed as a series of gates in a quantum computer through a process known as *Trotterisation*. Quantum computers, therefore, offer the possibility to efficiently encode more complete descriptions (i.e. ansätze) of electronic wavefunctions that are impractical for classical hardware.

Following the unitary coupled cluster approach, a number of ansätze amenable to quantum computers have been recently proposed [4], with different trade-offs between hardware efficiency and accuracy. Which one will perform better in real instances with molecules beyond classical capabilities is still unclear [4].

6.1.3 *Encoding electronic wavefunctions into qubits*

Once we have an efficient way of describing electronic wavefunctions in second-quantization, the next ingredient we need is to map them into qubit states. Because in all ansätze the wavefunction is provided by the action of combinations of creation and annihilation operators $\hat{a}_i^\dagger$ and $\hat{a}_i$ on Slater determinants, it is sufficient to map the action of fermionic operators to operations on qubits. The difficulty is to ensure that the fermionic rules in eq. (6.1.4) and eq. (6.1.11) are satisfied through the mapping.

One mapping that ensures such properties is the Jordan-Wigner transformation [7], where the occupation number $n_p \in \{0, 1\}$ of an orbital is directly stored into the state $|n_p\rangle \in \{|0\rangle, |1\rangle\}$ of a qubit. A single Slater determinant of a molecule with M orbitals can thus be efficiently encoded into a separable state of M qubits as

$$\overbrace{|n_1, n_2, \ldots, n_M\rangle}^{\text{Fock state}} \mapsto \overbrace{|n_1\rangle \otimes |n_2\rangle \otimes \cdots |n_M\rangle}^{\text{Qubits}}. \tag{6.1.21}$$

In order to ensure the fermionic rules in second quantization, i.e. eq. (6.1.11), the action of the creation and annihilation operators is mapped into the following qubit gates

$$\hat{a}_p^\dagger = \hat{\mathbb{1}}^{\otimes p-1} \otimes \hat{\sigma}_- \otimes \hat{\sigma}_z^{\otimes M-p} \tag{6.1.22}$$

$$\hat{a}_p = \hat{\mathbb{1}}^{\otimes p-1} \otimes \hat{\sigma}_+ \otimes \hat{\sigma}_z^{\otimes M-p}, \tag{6.1.23}$$

where $\hat{\sigma}_+ = |0\rangle\langle 1|$ and $\hat{\sigma}_- = |1\rangle\langle 0|$. Note in particular that, while the states are mapped locally, i.e. single Slater determinants correspond to separable states of qubits, the creation and annihilation operators, which represent local changes in the electronic system, correspond to highly non-local gates in the qubit framework.

The term $\hat{\sigma}_z^{\otimes M-p}$ implies that in this mapping the number of gates required to apply a single fermionic operator is $\mathcal{O}(M)$. The number of gates required in order to construct states satisfying the unitary coupled cluster ansatz (truncating at second excitation and using single step Trotterisation) via the Jordan-Wigner transform can be calculated to be $\mathcal{O}(M^3N^2)$ [4]. While in this thesis we will, for simplicity, focus on the Jordan-Wigner encoding, it is worth mentioning that further advantages in terms of resource requirements can be achieved when adopting other types of more complex mappings, e.g. the Bravyi-Kitaev encoding [4].

6.2 Quantum simulation based on quantum phase estimation

So far we have described how the electronic structure problem can be efficiently mapped into digital quantum hardware. In the next sections we will instead focus on describing

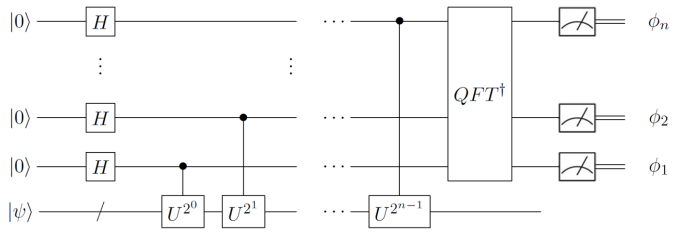


Figure 6.1: Circuit schematic for PEA, where the eigenphase is estimated with a n-bit resolution. A n qubits control register is initialised in the state $|+\rangle^{\otimes n}$, while the system register encodes the target state $|\psi\rangle$. After a series of control-U^{2^k} operations and an inverse quantum Fourier transform ($QFT^\dagger$), the first n bits of the eigenphase ϕ are obtained via measuring the control register in the computational basis.

the quantum algorithms that can be used to solve it in a scalable way (in principle) on a quantum computer. We start with algorithms based on quantum phase estimation.

6.2.1 *Quantum phase estimation*

The first protocol that promised to enable quantum machines to efficiently compute molecular energies was developed by Aspuru-Guzik et al. and is based on the quantum phase estimation algorithm (PEA) [8].

Quantum phase estimation is in general a fundamental sub-routine in quantum computing, and is necessary for harnessing many of its main applications, e.g the factorisation of large numbers [9, 10]. The goal of the algorithm is, given a unitary $\hat{U}$ and a quantum state $|\psi\rangle$, to learn an eigenvalue $e^{i2\pi\phi}$ ($\phi \in [0,1)$) of $\hat{U}$ within the support of $|\psi\rangle$. In particular, if the prepared state is close to an eigenstate $|\phi_0\rangle$ of $\hat{U}$, the algorithm returns with high probability the associated eigenvalue $e^{i2\pi\phi_0}$ up to an n-bit approximation. The schematic of the digital circuit for PEA is shown in Fig. 6.1.

To obtain an n-bit approximation of the eigenphase ϕ_0 (so that the error on the estimation is $\leq 2^{-n-1}$) in the PEA protocol we prepare an n-qubit control register, where each qubit is initialised in the state $|+\rangle$. The system register is instead initialised in a target state $|\psi\rangle$, which has non-zero overlap with the target eigenstate $|\phi_0\rangle$. We can expand the state $|\psi\rangle$ in an eigenbasis $\{|\phi_i\rangle\}$ of $\hat{U}$, i.e. $\hat{U}\,|\phi_i\rangle = e^{\mathrm{i}2\pi\phi_i}\,|\phi_i\rangle$, obtaining

$$|\psi\rangle = \sum_i c_i\,|\phi_i\rangle\,. \tag{6.2.1}$$

The initial separable state can thus be rewritten as

$$|+\rangle^{\otimes n} \otimes |\psi\rangle = \frac{1}{\sqrt{2^n}} \sum_i c_i (|0\rangle + |1\rangle)^{\otimes n} |\phi_i\rangle \tag{6.2.2}$$

$$= \frac{1}{\sqrt{2^n}} \sum_i \sum_{x=0}^{2^n-1} c_i |x_\mathrm{B}\rangle |\phi_i\rangle , \tag{6.2.3}$$

where x_B is a string of n bits given by the binary representation of x (padding zeros to the right if necessary). The algorithm then proceeds by applying a series of control-unitary operations, where, conditional on the state of the k-th qubit in the control register, the operation $U^{2^{k-1}}$ is applied on the system register. This produces the state

$$\frac{1}{\sqrt{2^n}} \sum_i \sum_{x=1}^{2^n-1} c_i |x_\mathrm{B}\rangle U^x |\phi_i\rangle = \frac{1}{\sqrt{2^n}} \sum_i \sum_{x=1}^{2^n-1} c_i e^{\mathrm{i}2\pi\phi_i x} |x_\mathrm{B}\rangle |\phi_i\rangle , \tag{6.2.4}$$

A crucial step for PEA is now to apply to the control register an inverse quantum Fourier transform gate $(QFT^\dagger)$[1], which acts as [10]

$$\frac{1}{\sqrt{2^n}} \sum_{k=0}^{2^n-1} e^{\mathrm{i}2\pi jk/2^n} |k\rangle \xrightarrow{QFT^\dagger} |j\rangle . \tag{6.2.5}$$

Assuming for simplicity that the phases $\phi_i \in [0, 1)$ assume a n-bit decimal representation

$$\phi = 0.\phi^{(1)}\phi^{(2)} \ldots \phi^{(n)} \equiv \frac{1}{2^1}\phi^{(1)} + \frac{1}{2^2}\phi^{(2)} + \ldots + \frac{1}{2^n}\phi^{(n)}, \tag{6.2.6}$$

with $\phi^{(j)} \in \{0, 1\}$, and defining

$$(2^n\phi_i)_\mathrm{B} = \phi_i^{(1)}\phi_i^{(2)} \ldots \phi_i^{(n)} \tag{6.2.7}$$

the bit representation of $2^n\phi_i$, we can write the resulting state as

$$(6.2.4) \xrightarrow{QFT^\dagger} \sum_i c_i |(2^n\phi_i)_\mathrm{B}\rangle |\phi_i\rangle = \sum_i c_i \left(|\phi_i^{(1)}\rangle \otimes \ldots \otimes |\phi_i^{(n)}\rangle\right) |\phi_i\rangle . \tag{6.2.8}$$

The n bits of the phase are now encoded in the n control qubits, and can then readout by measuring the control register in the computational basis. Upon measuring the eigenphase ϕ_i on the control registry, the system registry is collapsed to the associated eigenstate

1 Implementable on a n-qubit register via $\mathcal{O}(n^2)$ gates [10].

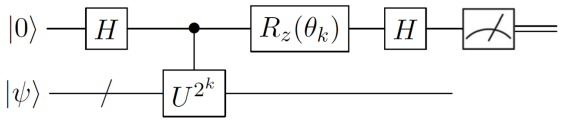


Figure 6.2: Circuit schematic for IPEA. In contrast to the standard PEA scheme, here only a single qubit is used in the control register. The algorithm iteratively learns the bits of the phase by adaptively choosing the parameter θ_k.

$|\phi_i\rangle$[2]. However, the eigenstate and the associated eigenphase on which we collapse is not determined a-priori, and in general we will collapse into $|\phi_i\rangle$ with a probability $|c_i|^2$. In particular, we have a probability $|c_0|^2$ to fall off into the target eigenstate $|\phi_0\rangle$. Moreover, if the phase ϕ_0 does not have an exact n-bit representation the probability of obtaining a correct eigenphase is further reduced to

$$p(\phi_0) = |c_0|^2 \frac{\sin^2(\pi\delta)}{2^{2n}\sin^2(\pi 2^{-n}\delta)}, \tag{6.2.9}$$

where $\delta = \phi_0 - \phi_0^{(n\text{-bit})}$ is the remainder of the true phase with the n-bit truncation [10, 11]. This probability can be rapidly increased by adding more qubits in the control register [10].

The PEA algorithm is then capable of efficiently calculating, with a good approximation and probability, eigenvalues of unitary operations. Note that, in order to have scalability, we require the control-U^k operations to be decomposed efficiently in gates acting on the system register. While this is not the case for general operations, we will see in section 6.2.3 that for the typical unitaries arising in quantum simulation applications such efficient decomposition can be possible.

6.2.2 *Iterative quantum phase estimation*

One drawback of the standard PEA approach is the requirement of many ancillary qubits for the control register. This demand makes PEA difficult to implement on near-term quantum hardware, where the number of available qubits is likely to be limited.

It turns out that all such control qubits are not really required. Phase estimation can, in fact, be implemented by replacing the control register with a single control qubit, and use classical processing of the data. The algorithm, called *iterative quantum phase estimation* algorithm (IPEA), was originally proposed by Kitaev [12] and further

2 Note that the system register is not measured, but is collapsed to an eigenstate by measuring only the control qubits. PEA can thus also be seen as a protocol to prepare eigenstates of a unitary (which eigenstate is prepared is heralded by the measured eigenphase), the success of which depends on the overlap on the initial state $|\psi\rangle$ with the target eigenstate.

elaborated, for instance, by Dobšíček et al. [11]. The general idea is to divide the task in multiple iterations of the experiments, each providing a single bit of the target phase. Such steps are iteratively repeated to infer all the bits starting from the least significant one and obtaining the most significant one from the last experiment.

The circuit schematic associated to each step of the algorithm is shown in Fig. 6.2. At each experiment, the system is initialised in the state $|\psi\rangle$ and the control qubit in $|+\rangle$. The system then undergoes a controlled-unitary operation. A rotation $R_z(\theta)$ along the $\hat{z}$ axis of the Bloch sphere by an angle θ is performed on the control qubit, which is finally measured in the computational basis after the action of a Hadamard gate. Suppose for simplicity that $|\psi\rangle = |\phi\rangle$ is an eigenstate of U, i.e. $U|\psi\rangle = e^{\mathrm{i}2\pi\phi}|\psi\rangle$, and that the associated eigenphase $\phi = 0.\phi^{(1)}\phi^{(2)}\dots\phi^{(n)}$ has a finite n-bit expansion[3]. Performing the controlled-U^{2^k} operation on the initial state $|+\rangle|\psi\rangle$ we obtain

$$\frac{1}{\sqrt{2}}\left(|0\rangle + e^{\mathrm{i}2\pi 2^k\phi}|1\rangle\right)|\psi\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle + e^{\mathrm{i}2\pi 0.\phi^{(k+1)}\phi^{(k+2)}\dots\phi^{(n)}}|1\rangle\right)|\psi\rangle\,, \tag{6.2.10}$$

where the bits $\phi^{(i)}$ with $i < k+1$ can be deleted as the phase is defined up to multiples of 2π. The state is separable so we can from now on neglect the system register. Performing the $R_z(\theta)$ rotation on the control qubit we obtain

$$\frac{1}{\sqrt{2}}\left[|0\rangle + e^{\mathrm{i}2\pi\left(0.\phi^{(k+1)}\dots\phi^{(n)}+\theta\right)}|1\rangle\right] \tag{6.2.11}$$

$$= \left[\frac{e^{\mathrm{i}\pi\left(0.\phi^{(k+1)}\dots\phi^{(n)}+\theta\right)} + e^{-\mathrm{i}\pi\left(0.\phi^{(k+1)}\dots\phi^{(n)}+\theta\right)}}{2}\right]|+\rangle$$

$$- \left[\frac{e^{\mathrm{i}\pi\left(0.\phi^{(k+1)}\dots\phi^{(n)}+\theta\right)} - e^{-\mathrm{i}\pi\left(0.\phi^{(k+1)}\dots\phi^{(n)}+\theta\right)}}{2}\right]|-\rangle \tag{6.2.12}$$

$$= \cos\left[\pi\left(0.\phi^{(k+1)}\dots\phi^{(n)}+\theta\right)\right]|+\rangle - i\sin\left[\pi\left(0.\phi^{(k+1)}\dots\phi^{(n)}+\theta\right)\right]|-\rangle \tag{6.2.13}$$

$$\xrightarrow{H} \cos\left[\pi\left(0.\phi^{(k+1)}\dots\phi^{(n)}+\theta\right)\right]|0\rangle - i\sin\left[\pi\left(0.\phi^{(k+1)}\dots\phi^{(n)}+\theta\right)\right]|1\rangle\,, \tag{6.2.14}$$

where we have neglected a global phase to obtain eq.(6.2.12), and applied the final Hadamard gate to obtain the bottom line. The probability to measure the control qubit in the state $|0\rangle$ is then

3 In case these assumptions are not satisfied, IPEA can probabilistically provide the correct n-bit truncation of the target eigenphase with the same probability as for the standard PEA approach, as given in eq.(6.2.9) [11].

$$p(0|k,\theta) = \cos^2\left[\pi\left(0.\phi^{(k+1)}\dots\phi^{(n)} + \theta\right)\right] \tag{6.2.15}$$

In order to infer the n bits of the phase with such circuit, we can design iterative steps as follows. At the j-th step ($1 \leq j \leq n$), we pick $k = n - j$. For $j = 1$, the angle of the rotation is chosen to be $\theta_1 = 0$, so that for the output probability we have

$$p_1(0) = \cos^2\left[\pi\ 0.\phi^{(n)}\right] = \begin{cases} 1 & \text{if } \phi^{(n)} = 0, \\ 0 & \text{if } \phi^{(n)} = 1. \end{cases} \tag{6.2.16}$$

The value of the last bit $\phi^{(n)}$ is then deterministically obtained from the the control qubit measurement. To infer the other bits we can iteratively define the value of the rotation angle at the j-th step (for $j > 1$) as $\theta_j = -0.0\phi^{(n-j+2)}\dots\phi^{(n)}$, where now we can assume $\phi^{(n-j+1)}\dots\phi^{(n)}$ to be known from previous iterations. In this way, the output probability for the control qubit at the j-th step is given by

$$p_1(0) = \cos^2\left[\pi\left(0.\phi^{(n-j+1)}\phi^{(n-j+2)}\dots\phi^{(n)} - 0.0\phi^{(n-j+2)}\dots\phi^{(n)}\right)\right] \tag{6.2.17}$$

$$= \cos^2\left[\pi\ 0.\phi^{(n-j+1)}\right] = \begin{cases} 1 & \text{if } \phi^{(n-j+1)} = 0, \\ 0 & \text{if } \phi^{(n-j+1)} = 1. \end{cases} \tag{6.2.18}$$

Repeating such iterations all bits of the phase can be deterministically calculated. As the precision on the phase estimation increases exponentially with the number of iterations, such task can be performed efficiently[4].

Note that if $|\psi\rangle$ is an eigenstate of the unitary, its state remains unchanged during the protocol, and can thus be recycled for all iterations of IPEA, if it is possible to reinitialise the control qubit after the measurement. As for PEA, if $|\psi\rangle$ is instead not an eigenstate, through IPEA the final state can be collapsed to a target eigenstate $|\phi_0\rangle$ with a probability $|\langle\psi|\phi_0\rangle|^2$, producing the associated eigenphase as well.

6.2.3 *Trotterization and calculation of Hamiltonian energies via phase estimation*

We can now describe an algorithm for the efficient exact calculation of electronic energies via PEA. The idea to start with is, quite simply, to map energies into phases. This can be done by noticing that, if $|\psi\rangle$ is an eigenstate of an Hamiltonian $\hat{H}$ with associated energy E, then it is also an eigenstate of the unitary evolution $U(\tau) = \exp(-i\hat{H}\tau)$:

4 Under the assumption that controlled-U^{2^k} operations can be efficiently implemented, as for standard PEA.

$$U(\tau)\left|\psi\right\rangle = e^{-i\hat{H}\tau}\left|\psi\right\rangle = e^{-iE\tau}\left|\psi\right\rangle = e^{i2\pi\phi}\left|\psi\right\rangle, \tag{6.2.19}$$

where the energy is now encoded[5] in the eigenphase ϕ via $E = -2\pi\phi/\tau$. A straightforward approach to calculate chemical energies is then to encode a target electronic wavefunction into the N-qubit state $|\psi\rangle$, for example using the Jordan-Wigner mapping discussed in section 6.1.3, and calculate E through the estimation of ϕ via PEA or IPEA.

As discussed above, in order to implement phase estimation in a scalable way we need to be able to efficiently implement controlled-$U(\tau)$ operations. This requires a decomposition of the controlled time evolution with a polynomial number of gates. One way of doing this is the Trotter decomposition [10, 14]. The procedure, called *Trotterization*, is a general prescription for approximating unitary time evolutions on a digital quantum computer. The idea is that, for most quantum systems, the Hamiltonian can be approximated by a sum of local interactions

$$\hat{H} = \sum_k \hat{h}_k, \tag{6.2.20}$$

where the local Hamiltonians $\hat{h}_k$ act on a subsystem with a constant number of elements. Acting on small subsystems, the size of which is independent of the total system size, the evolution under the local Hamiltonians $\exp(-i\hat{h}_k t)$ can be efficiently approximated on small scale digital quantum circuits. However, due to interactions between subsystems, the local Hamiltonians in general don't commute, meaning that we can't rewrite the global evolution as independent efficiently simulatable evolutions of smaller systems

$$e^{-i\hat{H}\tau} \neq \prod_k e^{-i\hat{h}_k\tau}. \tag{6.2.21}$$

Remarkably, such efficient approximation becomes possible if we subdivide the total evolution in many temporal steps, with the approximation being increasingly good for smaller time steps. This is mathematically enclosed in the Trotter-Suzuki approximation

$$e^{-i\hat{H}\tau} = \left(\prod_k e^{-i\hat{h}_k\tau/S}\right)^S + \mathcal{O}(1/S), \tag{6.2.22}$$

where S represents the number of Trotter steps [10]. The intuition behind this is to use a stop-frame approximation of the evolution, where each frame lasts for a time interval short enough so that the subsytems don't have time to interact. The approach requires

5 A subtle issue here comes from the fact that, as phases as defined up to 2π factors, we need to choose τ small enough so that $E\tau < 2\pi$ in order to avoid ambiguities in the energy estimation through PEA. In practice, this is typically done by posing a rough a priori upper-bound on the value of E [8, 13].

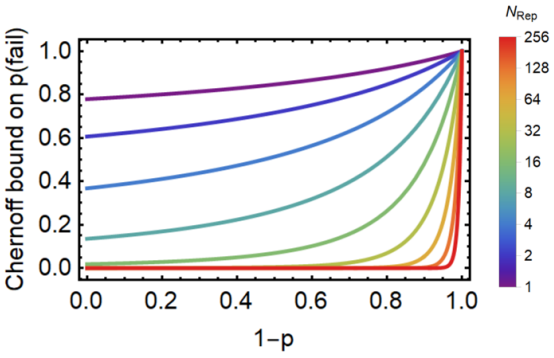


Figure 6.3: Chernoff Upper bound on error probability of failing in inferring a single bit when adopting majority voting in phase estimation. The bound is given as a function of error probability from the single-shot measurement $1-p$, and for various numbers of repetitions N_{Rep} in the majority voting scheme.

a polynomial number of gates, but in practice to reach high accuracy the circuit size can become challenging. For instance, to reach chemical accuracy (that is the accuracy typically required for practical chemical applications, usually assumed $\simeq$ 1 kCal/mol) using PEA-based simulations with approximately 30 qubits, $\mathcal{O}(10^{10})$ gate operations are needed in the Trotterization procedure to perform the exponentially long evolutions $U^{2^k} = \exp(-i\hat{H}2^k\tau)$ [15]. It is clear that avoiding the emergence of errors, due e.g. to decoherence or accumulation of gate imperfections, in circuits of such scale is quite a challenge, unless fault-tolerant quantum hardware is used. This represents an important drawback of PEA-based methods.

Additionally, in order to have the PEA algorithm collapsing to the correct eigenstate $|E_0\rangle$, the system needs to be initialised in a state $|\psi\rangle$ with a non-vanishing overlap with $|E_0\rangle$. Although this requirement is still a strong relaxation compared to finding the exact ground state, i.e. does not require highly accurate ansätze, it represents a non-trivial problem. In the case of ground states, a first possibility could be to initialise the system in the Hartree-Fock state. However, for large molecules the Hartree-Fock ansätz may have vanishing overlap with the true ground state. Many other approaches have then been proposed, including adiabatic state preparation [8], quantum cooling [16], and, as we will describe in sections 6.5 and 6.6, variational techniques [2, 17].

6.2.4 *Effect of noise on phase estimation*

A common argument against the PEA-based approaches to quantum chemistry, which we have already mentioned, is the low resilience to noise. To build up an intuition on the effect of errors on the algorithm, different noise models can be considered. As an example,

in Ref. [11] the authors examined the case where in the IPEA protocol the controlled-U operations are affected by a dephasing noise with decoherence time T_2, and errors on the control qubit are given by an additional rotation $R_x(\delta)$ with the angle δ having a Gaussian distribution with average $\tilde{\delta}$ and variance Λ_x^2. In this case, the probability of measuring the correct value of the k-th bit of the phase for single-shot measurements at the $(n-k+1)$-th step of IPEA is given by

$$p_k(\Delta_x, T_2) = \frac{1 + e^{-\Delta_x^2 - a2^k T_2} \cos\left(\pi 2^{k-n} \tilde{\delta}\right)}{2}, \tag{6.2.23}$$

where a is a system dependent parameter. In this equation the susceptibility to noise of IPEA is manifested, with the probability of inferring the correct bit decreasing exponentially with the relevant noise parameters.

In practical implementations, the success probability can be increased repeating each IPEA step multiple times and adopting a majority voting scheme, i.e. choosing the single-shot datum as the most frequent outcome. To investigate the improvement obtained with a majority voting scheme, we can use the Chernoff bound [10, 18], which provides an upper bound to the probability of failing to infer a bit by repeating the experiment N_{Rep} times:

$$p_{\text{MV}}(\text{fail}|p, N_{\text{Rep}}) \leq e^{-\frac{N_{\text{Rep}}(p-1/2)^2}{2p}}. \tag{6.2.24}$$

Here p is the probability of measuring the correct bit at each repetition, which, under the noise model used above, is given by eq. (6.2.23). The behaviour of the success probability $1 - p_{\text{MV}}(\text{fail})$ for different noise levels is plotted in Fig. 6.3. Although it is evident that adopting a majority voting approach an improvement in the noise robustness can be achieved, for high noise levels, which are expected in large-scale implementations on non-error-corrected quantum circuits, the failure probability is always dominant. In fact, while the majority voting scheme provides error resilience for small error rates, it can diverge rapidly once the error rate crosses a threshold, as we will experimentally show in section 6.4.

Moreover, if an error occurs, IPEA is not able to detect it, and thus has limited capability to correct it and reliably recover. Therefore, as noise become significant, errors accumulate and make the algorithm fail. This is particularly detrimental as the most significant digits are the last one to be estimated, and are thus the most likely ones to be affected from previous errors.

The susceptibility to noise of PEA-based quantum algorithms described above has induced a common belief that the approach requires fault-tolerant quantum hardware, and is thus not suitable for near-term devices [4, 17]. However, we will show in the next sections how novel techniques to phase estimation can now challenge this assumption.

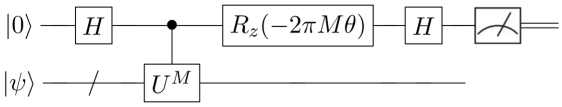


Figure 6.4: Circuit schematic for RFPE. The circuit consists of the same elements as for IPEA, and can thus be implemented on the same hardware. The difference between the two protocols is in the adaptive choice of the parameters M and θ: IPEA uses a fixed policy, while in RFPE they are adaptively chosen via Bayesian inference techniques.

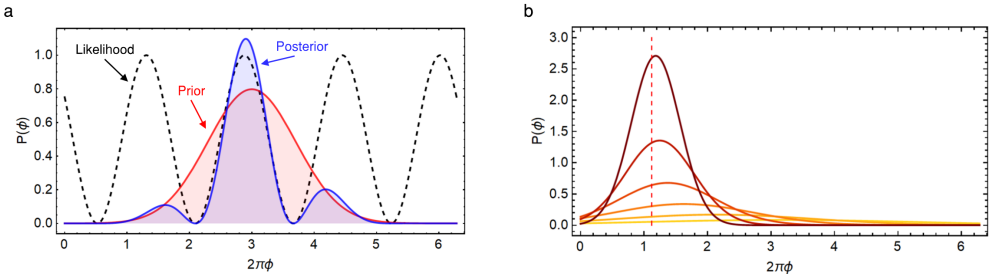


Figure 6.5: a, Representation of a Bayesian update of the prior distribution $P(\phi)$ (red) to the posterior $P(\phi|E)$ (blue), after a datum E with likelihood function $P(E|\phi)$ (dashed black) is measured. **b,** RFPE infers the unknown phase (dashed vertical red line) by repeating measurements and updating the prior distribution through Bayes' theorem. This allows us to collapse the prior distribution to the correct phase value.

6.3 A Bayesian approach to quantum phase estimation

In this section we will describe how techniques adapted from Bayesian inference applications can be used to enhance the phase estimation protocol. The algorithm, developed in 2016 by Wiebe and Granade [19], uses rejection filtering techniques to make Bayesian reasoning practical on a quantum computer, and was thus named rejection filtering phase estimation (RFPE). An experimental demonstration of the algorithm on a photonic chip will be reported in the next section.

6.3.1 *Bayesian quantum phase estimation*

The schematic of the circuit for RFPE is shown in Fig. 6.4. It can be observed that it is essentially identical to the one used in IPEA (Fig. 6.2), with the difference that now at each step the unitary exponent M and the rotation angle θ for $R_z(\cdot)$ are not chosen according to a fixed policy. In fact, while at a fundamental level RFPE involves the exact same experimental setups as standard IPEA, the main differences arise in the way RFPE chooses experiments and in how the experimental data is classically processed. In

particular, RFPE makes use of techniques initially developed for machine-learning tasks, in the form of Bayesian inference methods.

Bayesian inference is a formalism that aims to replicate the reasoning that leads to learning in intelligent beings. In this scenario, rather than describing the unknown eigenphase $2\pi\phi_0 \in [0, 2\pi)$ itself, we describe the experimentalist's knowledge of it. Such knowledge is encoded into the *prior distribution* $P(\phi)$, which is a fundamental object in Bayesian approaches. In RFPE, it models the algorithms confidence that a particular eigenphase ϕ corresponds to the eigenvector provided to the algorithm. As experiments are performed, the prior distribution is then updated based on the likelihood of the observed data. More in detail, if we perform a single run of the experiment in Fig. 6.4 using the parameters M and θ and an outcome $E \in \{0, 1\}$ is obtained from the measurement of the control qubit, the *likelihood function* for an arbitrary phase ϕ is given by:

$$P(E|\phi; M, \theta)/\kappa = \begin{cases} \cos^2(M\pi(x_i - \theta))/\kappa & E = 0 \\ \sin^2(M\pi(x_i - \theta))/\kappa & E = 1, \end{cases} \tag{6.3.1}$$

which is just a rewriting of eq.(6.2.15). Here, $\kappa \geq 1$ is a rescaling constant that can be used to slow down the inference (which may increase robustness to noise) [19]. The likelihood function $P(E|\phi)$ essentially represents the probability of obtaining the measured outcome E if ϕ was the true eigenphase of the system, and can be easily derived in the same way we did for the IPEA circuit in the previous section. Intuitively, if $P(E|\phi)$ is high then the experimenter's confidence on the phase value ϕ should increase once the datum E is measured, and should instead decrease if $P(E|\phi)$ is low. The update of the experimenter's knowledge, i.e. of the prior distribution $P(\phi)$, can be formalised using Bayes' theorem

$$P(\phi|E; M, \theta) = \frac{P(E|\phi; M, \theta)P(\phi)}{\int P(E|\phi; M, \theta)P(\phi)d\phi}, \tag{6.3.2}$$

where the updated prior distribution $P(\phi|E; M, \theta)$ is known as the *posterior distribution*. The update procedure in RFPE is depicted in Fig. 6.5a. Starting with a near-uniform prior over the $[0, 2\pi)$ interval, modeling no initial knowledge on the value of ϕ_0, the idea of the algorithm is then to proceed iteratively by repeating experiments and updating the prior distribution each time an outcome E is measured. At each step the posterior distribution is taken as the new prior for the successive step. In this way, as shown in Fig. 6.5b, the protocol is able to collapse the distribution to the correct eigenphase as more experience is gained from the experiment.

Some important differences with traditional phase estimation approaches can already be distinguished. First, rather than learning each of the phase bits individually, the RFPE algorithm gains information about every bit simultaneously. A consequence of this is that RFPE does not make hard decisions about bits at each experiment. Therefore, in

contrast to IPEA, errors are unlikely to be critical. In fact, RFPE is able to understand if a measurement result, coming from a noisy step, contrasts with previous assumptions, and deals with it simply widening the prior distribution, and keeps learning from subsequent experiments. RFPE learning ability is thus robust even when errors occur. As we will experimentally demonstrate in the next section, these properties significantly increase the noise resilience in phase estimation compared to traditional approaches [1, 19].

A second important difference is that at each step RFPE returns a posterior distribution over the phase rather than a single phase estimate. This means that not only we are provided with a value for the phase, e.g. given by the distribution mean, but we are also provided with an estimate of the algorithm's uncertainty in that value, e.g. the distribution standard deviation. This contrasts with standard PEA and IPEA, where the algorithm just returns a number with no error estimation. This capability is remarkable. In fact, providing correct and justified error regions is a central issue in estimation tasks.

A possible strategy for obtaining such uncertainty estimates from IPEA would be to determine the mean and the standard deviation of the measured eigenphase from repeated implementations of the full protocol. However, this approach introduces a significant slowdown in IPEA.

6.3.2 *Techniques for practical and efficient Bayesian inference*

We have described the general idea of the RFPE algorithm. We now proceed to describe some technical issues that need to be taken into account to make the inference process practical. A central problem when applying Bayes' theorem to update the prior distribution, as in eq.(6.3.2), is the estimation of the normalisation integral $\int P(E|\phi; M, \theta)P(\phi)d\phi$ over the interval $[0, 2\pi)$. The exact calculation of this quantity would in fact require the estimation of the likelihood for infinitely many values of ϕ, which is clearly not practical. The first solution that comes to mind is to discretise the $[0, 2\pi)$ interval with a sufficient number N of equally spaced points and describe the distributions only in these discrete phase values. In this way, we can approximate the integral in eq.(6.3.2) with a simple sum. However, using this discretisation the error in the phase estimation is at least π/N, that is the size of each bin in the discretisation. This means that, if we want to achieve a sufficient small error, the number of points required in the discretisation can be extremely large. As a consequence, even storing the prior distribution becomes quickly impossible. For example, if an error on the order of 10^{-9} is needed, then several gigabytes of memory will be needed to store the distribution, making such Bayesian inference impractical.

RFPE avoids this problem by adopting rejection filtering techniques [19]. The approach works by imposing a model for the prior distribution, which for convenience is here taken to be a Gaussian. The goal of each update is then to find the best model for the posterior distribution within the family of allowed models. The procedure, pictured in

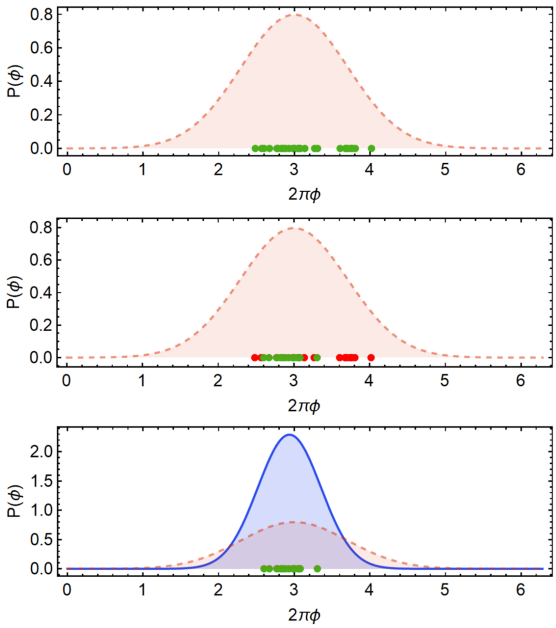


Figure 6.6: Functioning of the rejection filtering procedure for approximating Bayes updates. A fixed number of particles (points) are sampled from the prior (red dashed curve) Gaussian distribution (top panel). Once an outcome E is measured from the experiment, each particle is accepted with a probability equal to the associated likelihood $P(E|\phi)$ (central panel). The green particles in the central panel represent the accepted particles, while the red particles the excluded ones. The posterior distribution (blue curve) is given by the Gaussian with mean and covariance equal to those of the set of accepted particles (bottom panel), and used as new prior in the successive iteration.

Fig. 6.6, involves drawing a sample of N phase values $\{x_i\}$ $(0 \leq x_i < 2\pi)$, called *particles*, from the prior distribution and then computing the likelihoods $P(E|x_i)$. Each particle is then randomly accepted with probability equal to this likelihood $P(E|x_i)$. The mean and covariance matrix of the accepted particles can easily be shown to match the moments of the true posterior distribution [19]. The posterior distribution is thus taken as the Gaussian parametrised by the obtained mean and covariance, and is used as new prior for the successive iteration. Since the number of surviving particles decreases exponentially, fresh particles are added at each step, drawn from the new prior distribution. In this way, the number of points in the discretisation is kept constant, but their positions change during the inference to maintain a good description of the distribution. As a consequence, a limited number of particles does not pose limitations to the achievable precision on the phase estimation, as long as it is enough to provide a reliable description of Gaussian distributions. For RFPE, it is found that ~ 1000 particles are more than enough [1, 19]. Therefore, the rejection filtering approach allows the inference to be performed quickly in a memory limited environment without the need to compute costly functions.

An important remaining issue is that of choosing the parameters M and θ that, given the prior distribution, optimise the information gain from the experiment. A possible approach would be to choose the values of M and θ that minimise the posterior variance averaged over the possible outcomes of the experiment, namely the Bayes risk [19, 20].

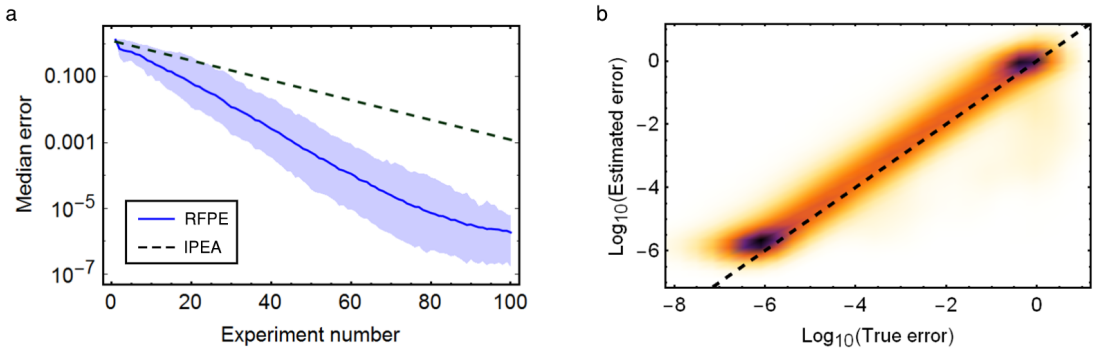


Figure 6.7: a, Simulated performance of RFPE. Blue line is the median error of the estimation averaged over 500 runs of RFPE. Shaded regions represent 95% confidence interval. Black dashed line represents the performance of standard IPEA when 10 repetitions are used to estimate the uncertainty. **b,** Smoothed histogram of the uncertainty estimated in the RFPE protocol, via the variance of the prior distribution, against the true median error in the estimation. Black dashed line would correspond to the ideal estimation; above this line we are overestimating the error, below we are underestimating the error.

However, the resulting optimisation can be too computationally expensive to carry out in online experiments. Nevertheless, optimisation heuristics can be investigated to make such calculation more practical. In particular, Granade et al. proposed a *particle guess heuristics* (PGH) which, for the class of likelihood functions employed in RFPE, can provide near-optimal experiments without requiring any precomputation [20]. In this approach, the experiment parameters are given by:

$$M = \lceil 1.25/\sigma \rceil, \qquad \theta \sim \mathcal{N}(\mu, \sigma). \tag{6.3.3}$$

Here $\lceil x \rceil$ indicates the first integer number greater than x, while μ and σ^2 represent the mean and variance of the prior distribution. $X \sim \mathcal{N}(\mu, \sigma)$ indicates a random variable sampled from the Gaussian distribution with mean μ and variance σ^2. The experiment parameters M and θ depend on the current estimate of the uncertainty, making the algorithm explicitly adaptive. In Ref. [19] the authors show that, under the assumption of a Gaussian prior, the PGH asymptotically saturates the Bayesian Cramer-Rao bound [10, 19, 20] (see also Fig. 6.8).

6.3.3 *Simulated performance of the RFPE algorithm*

The simulated performances of RFPE, considering an ideal implementation with no noise, are reported in Fig. 6.7 and Fig. 6.8. In all simulations 1000 particles were used in the PGH, and results are averaged over 500 independent runs of the RFPE algorithm with randomly chosen test phases. The exponential convergence of the algorithm to

the correct eigenphase of the system can be observed from Fig. 6.7a. In particular, the median error in the estimation decreases exponentially with the number of experiments performed in the iterative protocols. The capability of the algorithm to provide on-the-fly a reliable estimate of the error, provided by the variance of the prior distribution, is shown in Fig. 6.7b. It can be observed that the uncertainty reported from RFPE approximates very well the true estimation error with high probability. Typically, the estimated error overestimates the true error by a factor of ≈ 1.6, which is acceptable in practical circumstances.

The near-optimality of the PGH technique described before is studied in Fig. 6.8. Here, the phase estimation is considered as a metrology problem, where the resource is the number of times the unitary U is applied to the system $M_{\text{Tot}} = \sum_k M_k$, where M_k is the parameter M used in the k-th RFPE step. As this number is proportional to the total experimental time required to run the protocol, this quantity has a clear operational meaning. A fundamental limit in quantum mechanics for the asymptotic scaling of the error as a function of the resource used, i.e. M_{Tot} in our case, is given by the Heisenberg bound, which states that the error optimally scales as $1/M_{\text{Tot}}$ [19]. As can be observed from Fig. 6.8, the proposed PGH allows RFPE to asymptotically saturate the $1/M_{\text{Tot}}$ scaling, indicating near-optimal performance.

All these simulations considered no noise in the algorithm. However, the main advantage of the Bayesian approach over traditional methods comes from its expected improvement of the error resilience. This will be investigated, both numerically and experimentally, in the next section.

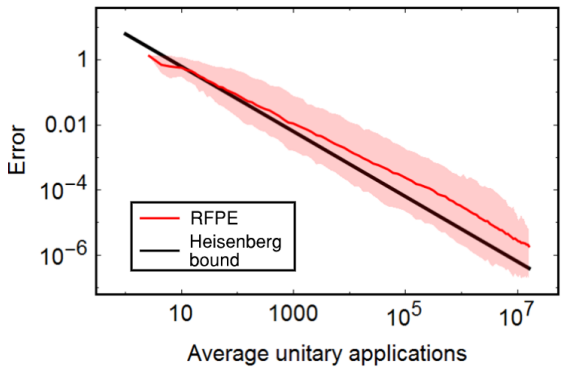


Figure 6.8: Performance of simulated RFPE as a function of total number of times M_{Tot} the unitary U is applied, which is proportional to the total experimental time. Red line is the error, averaged over 500 runs of RFPE. Shaded regions represent 95% confidence interval. Black line represents the fundamental limit in the asymptotic scaling given by the Heisenberg bound $\propto 1/M_{\text{Tot}}$.

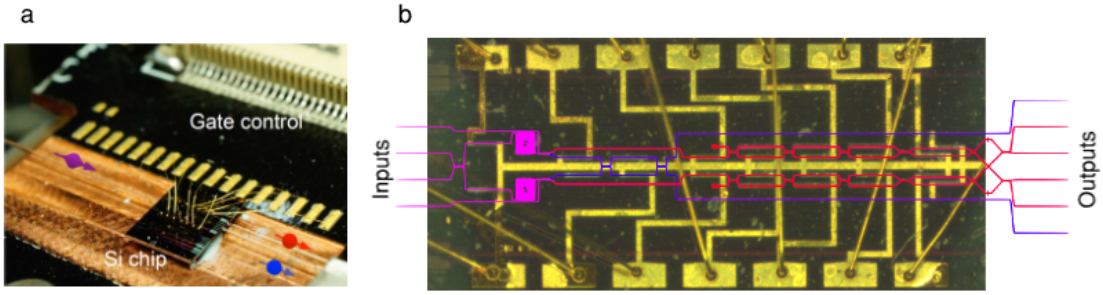


Figure 6.9: a, Photo of the silicon chip packaging. Pump light (purple) is horizontally coupled into the chip via a focusing lensed fibre and a spot-size converter. Similarly, signal and idler photons (red and blue, respectively) are collected out of the chip using two lensed fibres. The chip is glued to a PCB, to which it is electronically connected via wire-bonding. **b,** Top view of the silicon chip through an optical microscope. Waveguides associated to pump light, signal photon and idler photon are highlighted in purple, red and blue, respectively.

6.4 Experimental Bayesian quantum phase estimation on a photonic chip

In this section an implementation of the RFPE algorithm in a silicon photonic chip will be presented. Compared to the devices reported in chapters 3 and 4, the chip used in this experiment is more than 5 years older, and is far less complex. In fact, the remarkable evolution of the integrated quantum photonics field in the last few years is manifested in the difference between this device (which was state-of-the art only few years ago) and the current status of the technology. Nevertheless, the scheme used in the experiment contains some very interesting ideas in terms of photonic circuit designing. In particular, the scheme implemented allows us to perform arbitrary controlled-operations between two-qubits, providing the capability to perform quantum phase estimation protocols on-chip.

6.4.1 *Photonic device and experimental setup*

The silicon photonic chip and its packaging are pictured in Fig. 6.9. The device is a silicon-on-insulator chip fabricated using deep-UV photolithography and dry etching. In contrast to the devices reported in previous chapters, where light was vertically injected into the chip via grating couplers, here light is coupled horizontally into and out of the device via spot-size converters. As can be observed from 6.9a, a single lensed fibre was used to inject pump light into the chip, and two were used to collect the signal and idler photons generated and processed inside the chip. The input/output optical fibres were

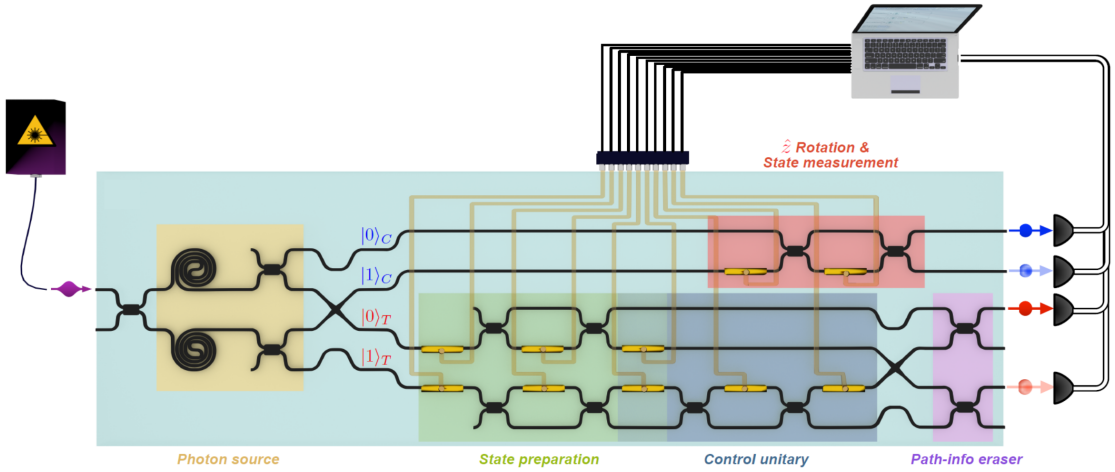


Figure 6.10: Schematic of the silicon quantum photonic device. The circuit can perform any controlled-U operation and any single-qubit state preparation and analysis. Photons are produced and guided in the silicon waveguides (black wires) and controlled by thermo-optical phase shifters. Coherent pump light was used to generate photons and off-chip superconducting nanowire detectors were used for the detection, both coupled to the chip through lensed-single mode fibres. The implementation of the algorithms was achieved by interfacing the quantum device with a classical CPU.

automatically recoupled before each measurement to maximise the coupling efficiency between waveguides and fibres. Approximately 8 dB coupling loss per facet was observed between the chip and a single-mode lensed fibre. The chip was mounted and wire-bonded on a PCB to electrically interface it and to maintain a stable chip temperature.

The on-chip operations on the pump light and on the photons are highlighted in Fig. 6.9b, with a schematic of the integrated circuit shown in Fig. 6.10. The integrated single-mode waveguides are designed with a width of 450 nm and thickness of 220 nm, and covered with a 1 μm silicon dioxide upper cladding. MMIs were used to realise integrated beam-splitters with near 50% reflectivity.

A schematic of the full experimental set up is pictured in Fig. 6.11. For the on-chip photon generation, a pumping scheme similar to the one adopted for the experiment in chapter 3 was used. A CW bright laser at 1551.9 nm was initially amplified off-chip by an erbium doped fibre amplifier. An off-chip pump power of $\approx$ 20 mW was used in typical measurement conditions. The spectral background was removed through the use of a WDM. A fibre-polarisation controller was used before the silicon device to maximise the coupling into the chip. Once coupled into the chip, the pump light is split via an MMI and used to coherently pump a pair of 1.2 cm long spiral sources. Equivalently to the scheme presented in chapter 3, the two coherently pumped sources generate a pair of

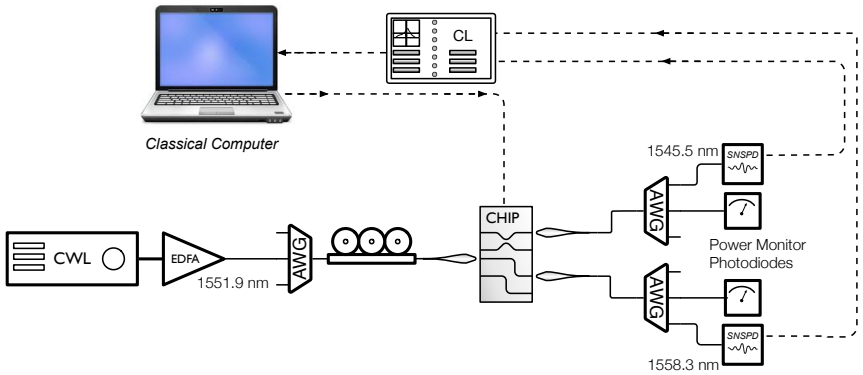


Figure 6.11: Schematic representation of the experimental set-up. The light from a CW laser source is amplified through an EDFA. The pump is filtered and the polarisation adjusted through a manual polarisation controller. The injection and collection of light in and out of the chip is done via lensed-fibres. The emerging light is then filtered to separate the spurious pump from the signal and idler photons, which are finally detected by SNSPDs.

path-entangled photons via the SFWM process, each encoding a qubit. The wavelengths of the signal and idler photons are selected at 1545.5 nm and 1558.3 nm, respectively. Considering the use in this experiment of non resonant SFWM photon pair sources and of a CW laser, the generation probability of multi-photon-pair events is negligible. The output photons, after evolving through the integrated device, were coupled into single-mode fibres. Off-chip filtering of the spurious pump light was achieved via two WDM (one per photon) with a 0.9 nm bandwidth and > 90 dB extinction. Photons were finally detected using SNSPD, with system detection efficiency of 80%, sub-100Hz dark counts, approximately 70 ps FWHM timing jitter and recovery times of approximately 50 ns. A maximal photon-pair rate of approximately 200 Hz was observed.

The active and reconfigurable quantum operations inside the silicon device rely on thermo-optical phase shifters. As reported in details in the next sections, the precision in controlling the implemented phases was characterised to be approximately 0.01 rad.

6.4.2 *Scheme to perform arbitrary controlled-unitary operations*

The logic circuits of the iterative phase estimation algorithms, both IPEA and RFPE (shown in Fig. 6.2 and Fig. 6.4), require controlled-unitary gates. As described in chapter 2, performing controlled operations is a notoriously difficult challenge in photonics, due to the lack of interactions between single photons. However, various approaches have been developed to implement such interactions via measurement-induced non-linearities [21–

24].

In our device, to implement arbitrary controlled-unitaries, we use an integrated entanglement-based approach introduced in Ref. [24]. The general idea is to induce non-linearities by enlarging the Hilbert space of the target photon and doing post-selection.

The circuit schematic is shown in Fig. 6.10. Coherently pumping the two waveguide SFWM sources, a pair of photons is generated in a superposition of being in the top source or in the bottom one. Using a Fock representation, the photon number state generated is written as $(|20\rangle + |02\rangle)/\sqrt{2}$ [25]. Two MMIs beam-splitters following the sources probabilistically split the photons. Post-selecting the cases where the signal photon is collected from the two upper modes and the idler from the four bottom modes the state becomes:

$$\frac{|1010\rangle + |0101\rangle}{\sqrt{2}}. \tag{6.4.1}$$

Note that, while in this chip the splitting is done probabilistically (with probability 1/4) via beam-splitters, it can also achieved deterministically using AMZI filters instead, as in chapter 3. To encode the qubit in the photons' path, for each photon we use the convention $|0\rangle \leftrightarrow |10\rangle$ and $|1\rangle \leftrightarrow |01\rangle$ to convert the number state to the logic state. The state obtained from the sources can thus be written as the maximally-entangled Bell state $(|0\rangle_C|0\rangle_P + |1\rangle_C|1\rangle_P)/\sqrt{2}$.

So far so good: we have generated maximally entangled states using the same approach as in chapter 3 for systems of dimension $d = 2$. Now comes the trick for implementing controlled operations.

Inserting an additional degree of freedom to the path of the signal photon, i.e. two ancillary modes, we obtain the entangled state $(|0\rangle_C|0\rangle_T|0\rangle_P + |1\rangle_C|1\rangle_T|1\rangle_P)/\sqrt{2}$. Note that what we have done is to add two ancillary modes to enlarge the Hilbert space of the signal photon, which now encodes a qudit of dimension $d = 4$ (ququart). This is equivalent to encode two qubits, which can now be manipulated singularly. The two components associated to the second qubit go through one R_z and one R_y gate to prepare the qubit in the same state $|\psi\rangle_T$ for both pairs of paths. These two components then go through two different operations: $\hat{I}$ for the upper pair of modes in the target register and $\hat{U}$ for the bottom ones, obtaining an entangled state of the form:

$$\frac{|0\rangle_C|\psi\rangle_T|0\rangle_P + |1\rangle_C(\hat{U}|\psi\rangle_T)|1\rangle_P}{\sqrt{2}}. \tag{6.4.2}$$

To obtain a superposition of these two different operations, we erase the path information between the two components of the target state. This is obtained employing two waveguide crossings and combining the modes in two MMI beam-splitters, giving

$$\frac{(|0\rangle_C|\psi\rangle_T + |1\rangle_C\hat{U}|\psi\rangle_T)|0\rangle_P + (|0\rangle_C|\psi\rangle_T - |1\rangle_C\hat{U}|\psi\rangle_T)|1\rangle_P}{2}. \tag{6.4.3}$$

Finally, projecting the third qubit into $|0\rangle_P$, the final state can be represented as

$$\frac{|0\rangle_C|\psi\rangle_T + |1\rangle_C\hat{U}|\psi\rangle_T}{\sqrt{2}} \tag{6.4.4}$$

which is equivalent to applying the desired arbitrary control-unitary operation. The circuit required for RFPE and IPEA (Fig. 6.2 and Fig. 6.4) is then finalised by implementing an arbitrary local operation on the control photon via an MZI and a phase shifter (see Fig. 6.10).

The scheme can be generalised to an an arbitrary number of photons in the system register [24]. However, being a post-selected scheme, the approach is not scalable. In particular, the probability of post-selecting the desired cases decreases exponentially with the number of qubits (as 2^{-n}). Nevertheless, it represents an interesting circuit which can enable a wide range of small-size experiments previously inaccessible to integrated quantum photonics.

Note also that the approach can also enable us to cascade many operations on the target qubit, which corresponds to multiple applications of controlled-unitary operation. This could be useful, for instance, to apply different Trotter steps in a Trotterisation procedure for quantum simulation. In our particular implementation, however, the circuit implements a single operation, meaning that in all the experiments we will consider a simplified case with a single Trotter step.

6.4.3 *Characterisation of phase errors in the integrated photonic device*

In the experiment we will focus on precision and robustness estimation for different types of algorithms. In order to perform such investigation, we first need to have a reliable model to describe the implemented operations and a characterisation of the noise level in the photonic device.

In our set-up we adopt on-chip thermo-optical phase shifters to implement both the quantum state preparation and unitary evolution of the qubits. In two-photon integrated optics experiments the main source of error comes from noise in controlling these phases. In more detail, the phase shifters consist of metallic thin-film tracks on top of the silicon waveguides, as described in section 2.3.1. When supplied with DC current, the phase shifters act as non-Ohmic resistive heaters, thus changing locally the refractive index of the waveguide. This change induces a phase shift, which is linear with the dissipated electrical power in the phase shifter. Errors in controlling these phases can arise mainly from three types of noise: finite resolution of the electrical power dissipated, electrical cross-talk (e.g. due to non-zero resistance of shared grounds), and thermal cross-talk.

Because all phase shifters in the chip were connected to a single electrical ground

(see Fig. 6.5b), electrical cross-talk was a major source of noise in this experiment [6]. To remove such noise, each of the heaters was driven and controlled independently using multi-channel current drivers. In fact, if voltage driving was adopted, then the power actually dissipated on a single phase-shifter would actually be coupled to the voltages applied to all other heaters due to the non-zero resistance of the common ground. By driving directly the current injected in each heater, rather than the voltage, such dependence is avoided, and the electrical cross-talk problem resolved.

Each current driver was hosted in a board featuring 12 bits DAC, controlled by the experiment software via a RS-232 interface, allowing us to have high resolution on the current applied, hence the electrical power dissipated.

Having adopted such precautions, the main residual noise in the experiment was thermal cross-talk between the heaters. This type of noise arises when thermally induced unwanted changes in the refractive index of other optical paths (due to heat flowing across the dioxide layer and the electrical connections) alter the implemented phase shifts in a non-controllable way, introducing systematic errors in the setup.

In order to compensate for this effect, different calibrations, using classical light, for different chip configurations were preliminarily run. For each configuration, all the heaters were constantly driven at a certain current, set to implement an appropriate state preparation and/or gate operation, except the heater to be calibrated. This heater is supplied with a range of currents (and thus electrical power, P_{el}), and oscillations in optical power P_{op} at the output of the corresponding optical path are recorded. The collected data can be fitted according to the non-linear function

$$P_{op} = B + A\cos\left[\frac{2\pi}{T}(P_{el} - P_{\Phi})\right] \tag{6.4.5}$$

where B is a background, A is the maximum amplitude of the signal, T its period, and P_{Φ} is the offset power value for the heater in such a configuration. These parameters can be obtained for each phase shifter by fitting the model to the measured values of P_{op} as a function of P_{el}. Detailed calibrations are reported in Fig. 6.12 for three exemplary phase shifters, where the relevant statistical parameters of each fit are also reported. All fits show R^2 values close to one, thus suggesting that the model can adequately reproduce the data observed from the measurements. Also, the t-statistic and p-value associated to each parameter, testing its statistical significance in the model, are reported. In most cases, high t-statistics and low p-values are observed, with the exception of the background parameter B in those cases where the fringe visibility is particularly high (i.e. no significant background is present). This data analysis gives evidence of the suitability of the model in Eq. 6.4.5 to describe the operations in our device.

6 For the devices reported in previous chapters of this thesis the electrical cross-talk was not present as each heater had an individual ground. Voltage driving was in fact used in chapters 3 and 4.

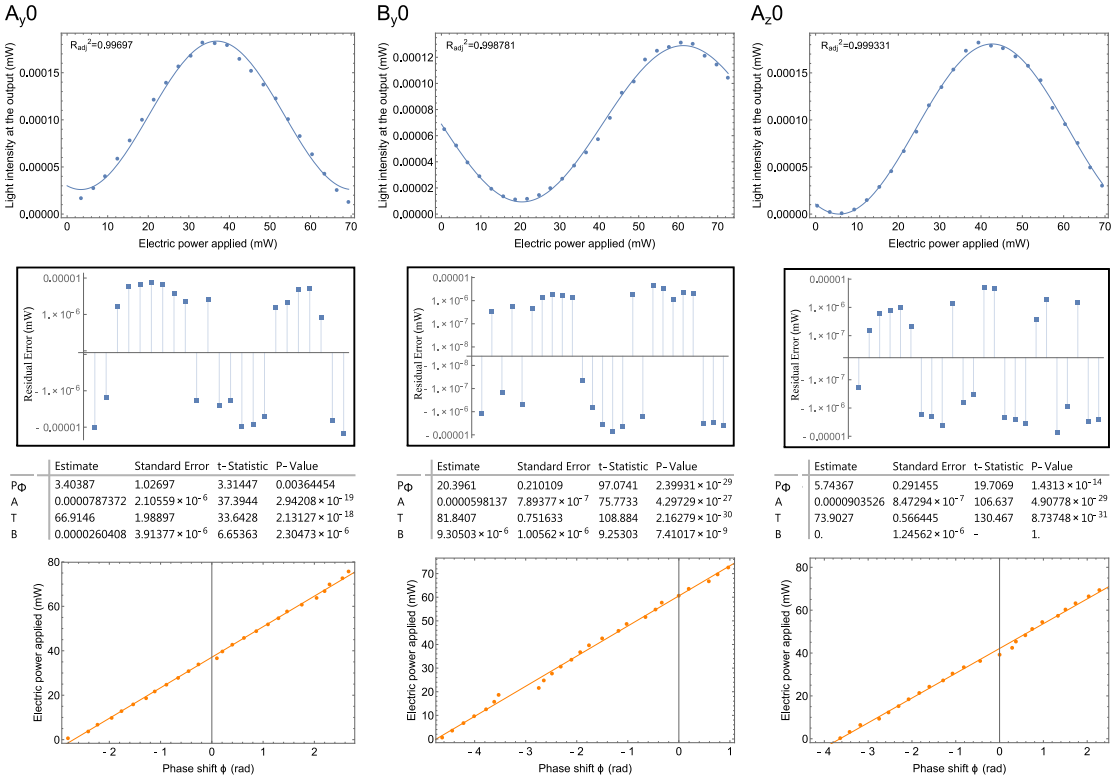


Figure 6.12: Characterisation of the optical phase control for three exemplary phase shifters in the device, chosen as a worst case (A_y0), an average case (B_y0) and a best case (A_z0). For each phase shifter it is reported in column, from top to bottom: i) the nonlinear fit of raw classical characterisation data while measuring the optical power output at the corresponding optical port. The fitting is performed using the nonlinear Eq. 6.4.5. The value of the R^2 fit parameter is reported in the plot for each separate case. ii) A plot of the residuals corresponding to the nonlinear fits. iii) A table of the relevant statistical parameters obtained in the nonlinear fit: the fitted value, its statistical uncertainty, and tests of its significance (t-statistic and p-value). iv) Linear dependence of the phase shift implemented by each heater against the corresponding electrical power applied.

Given the model, an estimation of the phase noise level can also be obtained from analysing these fits. In particular, after the calibration, the targeted phase $\bar{\varphi}$ for a heater is obtained driving the heater with P_{el} such that:

$$\bar{\varphi} = \frac{2\pi}{T}\left(P_{el} - P_{\Phi}\right) \tag{6.4.6}$$

and therefore, propagation of stochastic errors in P_{el}, T, P_{Φ} affect $\bar{\varphi}$. The relative statistical uncertainties obtained from the non-linear fits are $\sigma_{P_{\Phi}}/\max(P_el) \simeq 0.2\%$

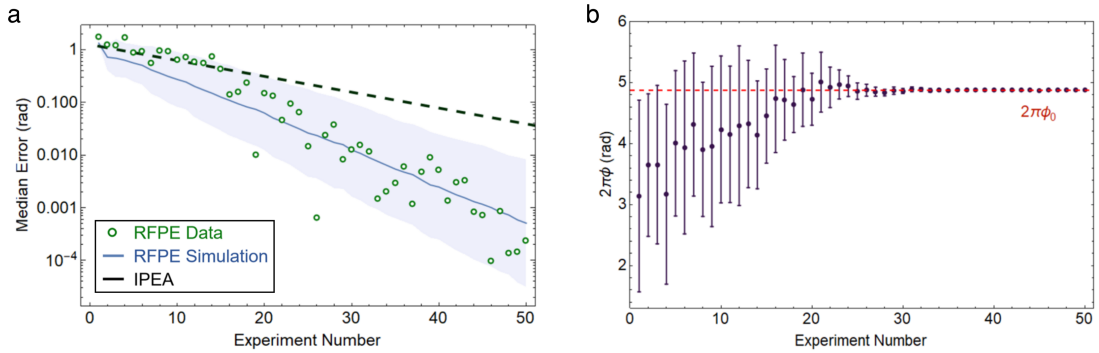


Figure 6.13: a, Evolution of the median error in the phase estimation for RFPE. The error is evaluated with respect to the true phase value $2\pi\phi_0 = 4.8741$ rad (related to the energy of the dissociated H_2 molecule). The initial prior distribution is $\mathcal{N}(\pi, \pi^2)$. Data points show an exponential shrink of the error, compatible with simulations (blue line) of the device performance averaged over 1000 runs of the RFPE algorithm (shaded area: 67.5% credible interval). The dashed black line denotes the convergence of the standard IPEA where 10 repetitions are used to estimate the statistical uncertainty. **b,** Convergence of the prior distribution in the RFPE algorithm to ϕ_0 (red line). Each point and its associated error represent the median and standard deviation of the prior distribution at different steps of the algorithm.

and $\sigma_T/T \simeq 1.1\%$. Inaccuracies in the current supplied by the driver to each heater ($\pm 0.005\ mA$), affecting the actual value of P_{el}, can be neglected as, using drivers with a 12 bit resolution, they are typically less than 0.04% for all the heaters. Confining the propagation of errors to only T, P_Φ, and averaging over the full interval of P_{el} adopted in the experiment (approximately $5 - 80$ mW, slightly different for each heater due fluctuations in the resistance values), an average precision of $\sigma_{exp} \simeq 0.01$ rad can be estimated for the experimentally implemented phases.

6.4.4 *Implementation of phase estimation algorithms*

Having described the photonic scheme able to implement the circuits required for RFPE and IPEA, and having built up a reliable model for the on-chip operations, we can now proceed to report the actual experimental results.

The rapid reconfigurability and the high precision of the silicon photonic device allowed us to correctly implement a full RFPE and IPEA algorithm in a typical time interval of few minutes, where most ($\gtrsim 95\%$) of the time was spent in photon collection rather than reconfiguring the circuit. As is usual in photonic implementations [2, 13, 26], in both RFPE and IPEA the value of the single-shot data E for each step of the algorithms was determined via majority voting, where approximately 2000 photons were collected per measurement to have enough statistics.

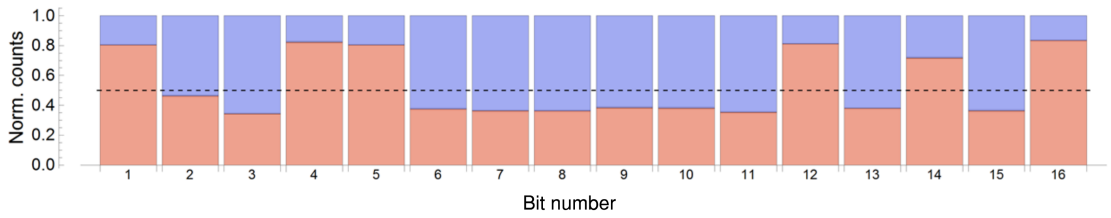


Figure 6.14: Example of experimental 16-bit phase reconstruction via IPEA on the silicon device. For each bit, red (blue) bars report the relative number of events the outcome was measured to be 1 (0). The dashed line represents the 0.5 threshold, above which majority voting assigns the bit to be 1, and 0 otherwise. In the reported example, the binary 16-bit representation of the true phase tested was 0.1001100000010101, which is correctly reconstructed by the algorithm.

For the implementation of RFPE, the PGH described in previous sections was used. For the stochastic representation of the prior, 1000 particles was found to be reliable. The initial prior distribution was set to a Gaussian $\mathcal{N}(\pi, \pi^2)$, which approximates a uniform prior over the $[0, 2\pi)$ interval.

Typical results for a single experimental run of RFPE are shown in Fig. 6.13. As a test, in this exemplary case the system qubit is prepared in the ground-state of a molecular Hamiltonian for the H_2 molecule, described in the next section. The behaviour reported in Fig. 6.13a experimentally demonstrates that the estimation converges exponentially to the true eigenphase $2\pi\phi_0$. For this case an error as low as $2.4 \cdot 10^{-4}$ rad is achieved within 50 experiments on the quantum device. This error is in good agreement with the standard deviation of the final posterior Gaussian distribution ($\simeq 4.2 \cdot 10^{-4}$ rad after 50 experimental steps), whose evolution in shown in Fig. 6.13b. This confirms experimentally that the algorithm is able to provide a reliable uncertainty estimate.

As the circuit required is the same, also IPEA is tested in the same chip. The bits of the phase experimentally reconstructed in a typical 16-bit IPEA run are reported in Fig. 6.14. As already discussed in previous sections, a single run of IPEA does not provide a rigorous statistical uncertainty on the phase value calculated with the algorithm. However, a possible frequentist strategy is to repeat the phase estimation by running IPEA multiple times and taking the standard deviation of the measured phases. In the experiment, we will use 10 repetitions of IPEA to estimate a statistical uncertainty on the phase. Note that in this configuration RFPE outperforms IPEA in terms of efficiency, as is experimentally demonstrated in Fig 6.13a.

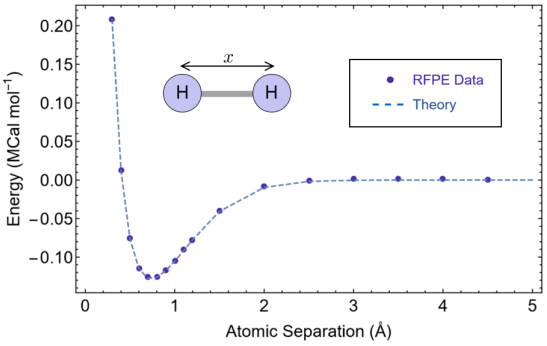


Figure 6.15: Bonding energies of the H_2 molecule for various atomic distances (x in the inset), calculated using RFPE with 50 steps and 1000 particles. Uncertainties, within chemical accuracy, are smaller than the markers and neglected in the plot for more clarity. The dashed line represents theoretical ground energies. Inset: schematic of the H_2 molecule.

6.4.5 *Photonic calculation of molecular energies via RFPE*

We have now developed all the preliminary ingredients required to experimentally implement simple quantum chemistry calculations on our small-scale digital photonic quantum computer. The molecular system we are going to test is quite simple: the H_2 molecule (see inset of Fig. 6.15) [13, 27]. The eigenstates of the molecular Hamiltonian, for different values of the atomic separation a, are mapped into the qubit basis using the Jordan-Wigner transformation, described in section 6.1. In a minimal basis, the molecular Hamiltonian can be block diagonalised in two 2×2 matrices [13]. A single qubit in the control register is then enough to encode the eigenstates in the 2-dimensional subspaces. Preparing the target qubit in the molecular ground-state, we can then use phase estimation to calculate the ground energies for the Hamiltonian corresponding to different atomic distances between the two Hydrogen nuclei, reconstructing the potential energy landscape for the H_2 molecule.

A difference with previous implementations of PEA-based digital quantum simulations [13, 27] is the use of the Bayesian approach to phase estimation. In particular, our experiment represents a benchmark to test how RFPE performs in these applications. The experimental results are shown in Fig. 6.15. The energies estimated via RFPE match the theoretical values with high precision. The average median error in the energy estimation for the reported dataset is 0.72 kCal/mol, providing a precision higher than chemical accuracy ($\simeq 1$ kCal/mol) [13, 27].

6.4.6 *Testing the robustness of RFPE to experimental noise*

As discussed in previous sections, the main advantage of the Bayesian approach over traditional methods comes from its expected reliability on non-fault-tolerant devices. While such expectations were born out in simulation [19], they had not not been verified in practice. In this section we report such verification. This is possible via the capability

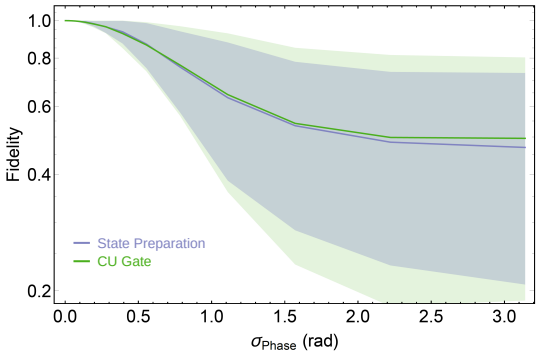


Figure 6.16: Simulation of the silicon photonic device performance in terms of the average state preparation fidelity (violet) and controlled gate fidelity (green) for different levels of Gaussian noise in the phase shifters, as described by σ_{phase}. Shaded areas represent 67.5% confidence intervals.

of implementing both IPEA and RFPE algorithms on the same device, allowing us to set a unique context for a fair comparison. We experimentally investigate the resilience of both protocols to two sources of noises: gate errors and decoherence.

The infidelity of unitary operations is a well known problem existing in quantum hardware, and is typically given by a noisy control and imperfect manufacture and calibration of the components. On integrated quantum photonic devices, this type of noise is mainly due to control noise and residual cross-talk of the phase gates. The characterisation for this type of noise in our device was reported in previous sections, where a precise control of the thermo-optical phase gates was demonstrated, with errors of approximately 0.01 rad. Such high controllability allows us to add a tunable level of noise on the phase gates to study the robustness of the protocols. This is achieved by replacing the correct phases $\bar{\varphi}$ required to implement the unitary transformations with synthetic values φ sampled from a Gaussian distribution $\varphi \sim \mathcal{N}(\bar{\varphi}, \sigma_{\text{phase}})$. The parameter σ_{phase} mimics a Gaussian noise in the phases, which in turn introduces a controllable noise in the fidelity of both the implemented state preparation and the unitary evolution. This effect can be quantitatively evaluated using the calibration model in eq. (6.4.5). The obtained infidelities as a function of σ_{phase} are shown in Fig. 6.16.

Fig. 6.17a shows the convergence of both RFPE and IPEA scanning σ_{phase} up to 0.55 rad, corresponding to average state fidelity 94% and gate fidelity 91% expected in the chip. We report the performance of RFPE with 100 steps, compared to a 16-bit IPEA, averaged over 10 runs to estimate the error bars, i.e. 160 total experiments per data point. Since a reasonable error bar estimation requires a higher number of experiments for IPEA than RFPE, the relative rates of convergence are not immediately obvious from these figures. Rather, these plots illustrate how the performance of each algorithm is affected by increasing errors, to compare their robustness to noise.[7]

7 Note also that the uncertainties in Fig. 6.17 are given by Bayesian credible interval for the RFPE case, and by frequentist credible intervals for IPEA. A direct comparison of the uncertainty bounds is then

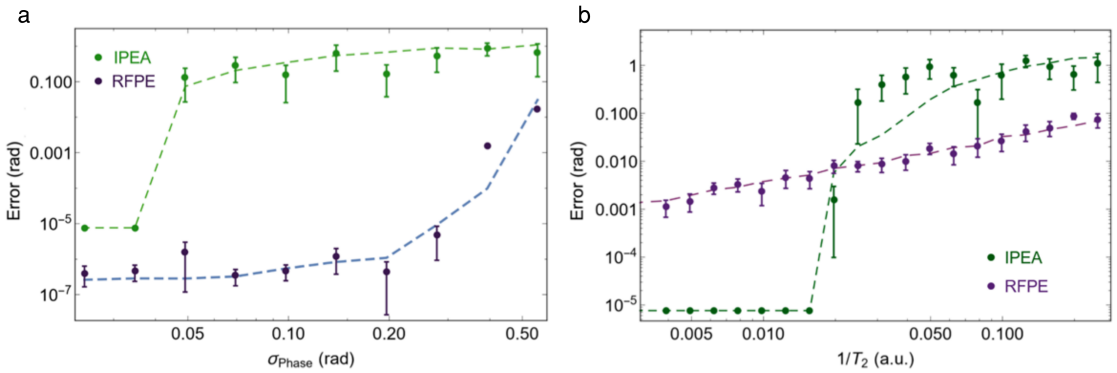


Figure 6.17: Effects of different experimental noises on phase estimation strategies. **a,** Infidelity of quantum operation. Each of the correct phases $\bar{\varphi}_i$ for the phase gates is synthetically replaced with a Gaussian distributed $\varphi_i \sim \mathcal{N}(\bar{\varphi}_i, \sigma_{Phase})$, where σ_{phase} represents a noise in the phases. **b,** Decoherence. For IPEA data, experiments were repeated 10 times with 16-bit accuracy to evaluate median error and error bars, while the RFPE data were collected from a single run, after 100 measurements, and directly used to evaluate the error and uncertainty within the algorithm. Error bars for the estimated phase represent in both plots either a 67.5% credible region for RFPE, either a 67.5% confidence interval for IPEA. In the cases where error bars are smaller than the markers they have been omitted for clarity. Points are experimental data and dashed lines are simulations averaged over 1000 runs. The simulations take into account the characterised residual phase noise in the device.

For $\sigma_{\text{phase}} \geq 0.05$ rad IPEA dramatically decreases in accuracy and becomes quickly unreliable. As described in previous sections, this occurs because while the majority voting scheme provides error resilience for small error rates, it can diverge rapidly once the error rate crosses a threshold. On the other hand, in this regime the performance of RFPE is initially only slightly affected, maintaining a very high level of precision even when IPEA fails. In order for RFPE to be substantially affected we require σ_{phase} higher than 0.3 rad, a value much higher than the actual experimental noise in our device.

The second type of noise investigated is decoherence. While being an important limitation in many quantum computing experiments, decoherence plays a minor role in integrated quantum photonics due to the intrinsic stability of the platform. It must then be introduced artificially in our experiment. In order to simulate it, coincidence counts provided by the SNSPDs for the $|0\rangle$ and $|1\rangle$ projectors on the control qubit were progressively flattened out by classical post-processing and combined with Poissonian noise in the measurements. In agreement with the depolarizing noise model, the presence

impossible except in the asymptotic limit. Nevertheless, the errors in the phase estimation, that is the primary point of comparison, are clearly comparable.

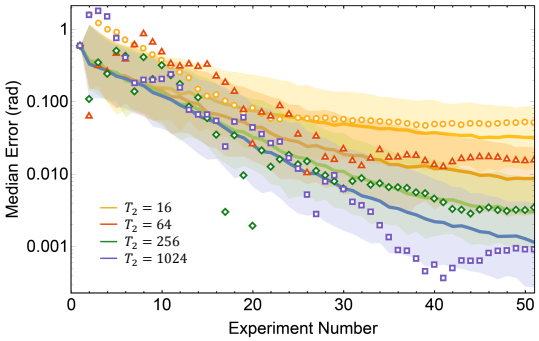


Figure 6.18: Experimental data (points) and simulation (lines) of the behaviour of RFPE under different decoherence times T_2. Shaded areas represent a 67.5% credible intervals from data simulated averaging over 500 runs of RFPE.

of a normalised decoherence time T_2 was mimicked via:

$$P'(D|\phi) = e^{-M/T_2} P(D|\phi) + \frac{1 - e^{-M/T_2}}{2},$$

where $P(D|\phi)$ is the statistics obtained from the photon coincidence counts for the outcome $D \in \{0, 1\}$. We introduce this noise model while processing online the output data during the iterative process, thus affecting the choice of the experiments. In this way it is possible to simulate the behaviour of RFPE and IPEA in systems that are prone to this model of decoherence.

In Fig. 6.17b the action of the depolarising noise, up to $T_2 = 4$, is reported for both IPEA and RFPE. The performance of IPEA has a substantial and sharp deterioration at T_2 close to 32, whereas the median error of 100-step RFPE decreases only polynomially with $1/T_2$, maintaining an error $\mathcal{O}(10^{-2})$ even in the regime where conventional IPEA fails to provide any reliable estimate of the phase.

In the presence of characterised depolarising noise an optimised value for M is given by $\min(\lceil 1.25/\sigma \rceil, T_2)$ [19], which, however, implies that, when decoherence is significant, the performance of RFPE degrades significantly. This behaviour is exhibited by the experimental data in Fig. 6.18, where the convergence of RFPE is reported under the action of various T_2. We observe that RFPE ceases to learn exponentially quickly when $1/\sigma \sim T_2$, after which the algorithm continues to learn at a polynomial rate, unlike IPEA [19].

6.4.7 *Experiment discussion*

The capabilities of the silicon photonic chip to integrate the reconfigurable circuit required for both IPEA and RFPE with high precision was key to perform the implemented tests. The main result of the experiments was the experimental demonstration of a remarkable higher resilience to noise of the Bayesian approach to phase estimation, which is instead

a significant practical weakness for the standard PEA and IPEA methods. While RFPE was here demonstrated on a silicon photonics platform, the approach is also clearly amenable to be efficiently performed on any scalable quantum architecture. Given the importance of phase estimation as a fundamental subroutine in many quantum protocols, from quantum metrology to Shor's algorithm, our result has a potential impact on a wide range of of applications on pre-fault-tolerant quantum hardware. For instance, in collaboration with Prof. Fedor Jelezko's group in Ulm, we have recently shown that, by using a small modification in the Bayesian inference methodology, the approach can provide great sensitivity advantages for room-temperature quantum magnetometers based on NV-centres in diamond [28].

Of course, digital quantum simulation is another application that can greatly benefit from the RFPE algorithm. However, some significant challenges for phase estimation-based quantum simulation still remain. In particular, the need of performing a large series on unitary operations on the state prepared, needed for the controlled-U^M gate, significantly increases the depth of the circuit and requires long coherence times to achieve a good precision in the estimation. This requirement remains very challenging for current quantum platforms.

6.5 Variational quantum eigensolvers

In this section we will describe a different method to approach the electronic structure problem that does not require quantum phase estimation. The algorithm, generally referred to as *variational quantum eigensolver* (VQE), represents a hybrid quantum-classical approach that combines quantum hardware with well-developed classical methods for quantum chemistry. While the number of operations that need to be performed on the quantum hardware is significantly reduced compared to PEA-based approaches, which is important for near-term noisy experiments, VQE strongly depends on the reliability of the ansätze to accurately describe the electronic molecular states of interest.

6.5.1 *Hybrid quantum-classical algorithm*

Consider again mapping an electronic structure to a Hamiltonian $\hat{H}$ acting on M qubits, e.g. via the Jordan-Wigner transformation, with associated ground state energy E_0. If the system is prepared in a trial wavefunction $|\psi(\boldsymbol{\theta})\rangle$, described within a given ansatz by a set of parameters $\boldsymbol{\theta}$, then the variational principle famously states that[8] the expectation

8 Note that, due to linearity of quantum mechanics, the variational principle holds also if the trial state is not pure: $\mathrm{Tr}\left[\rho(\boldsymbol{\theta})\hat{H}\right] \geq E_0$. This allows us to make use of it even in real situations where experimental noise produces non-ideal states.

value of the Hamiltonian can not be less than the ground energy:

$$E(\boldsymbol{\theta}) = \langle\psi(\boldsymbol{\theta})|\,\hat{H}\,|\psi(\boldsymbol{\theta})\rangle \geq E_0. \tag{6.5.1}$$

This implies that the ground state energy can be approximated, up to the suitability of the ansatz, by finding the parameter values $\bar{\boldsymbol{\theta}}$ that minimise the expectation value $E(\bar{\boldsymbol{\theta}})$.

A wide class of classical methods are based on this idea, where the electronic configuration is approximately represented in a CPU by an efficient ansatz $|\psi(\boldsymbol{\theta})\rangle$, and an optimisation algorithm is used to variationally drive the parameters into the ground state by minimising the objective function $E(\bar{\boldsymbol{\theta}})$. However, there are two sub-routines that make this approach inefficient in classical computers.

First, as we have seen in section 6.1.2, existing efficient classical ansätze are limited and can not provide a reliable description for certain classes of chemical systems. More complete ansätze, e.g. the unitary coupled cluster approach described in 6.1.2, are known, but cannot be efficiently implemented on classical machines. Secondly, classical computers are in general unable to efficiently evaluate the expectation value $\langle\psi(\boldsymbol{\theta})|\,\hat{H}\,|\psi(\boldsymbol{\theta})\rangle$.

The idea of the VQE algorithm is to use quantum hardware to perform these two subroutines: ansatz state preparation, and measurement of the expectation value. A classical computer is then used to update the parameters using an optimisation algorithm. The protocol is schematised in Fig. 6.19. The trial states for a given ansatz are prepared in the quantum circuit by generating an initial reference state $|\psi_{ref}\rangle = U_{ref}\,|\mathbf{0}\rangle$, representing for example a Hartree-Fock wavefunction, and then applying a set of unitary gates $U(\boldsymbol{\theta})$ parametrised by the values $\boldsymbol{\theta}$ to generate the ansatz state

$$|\psi(\boldsymbol{\theta})\rangle = U(\boldsymbol{\theta})\,|\psi_{ref}\rangle\,. \tag{6.5.2}$$

For example, in the case of the unitary coupled cluster ansatz discussed in section 6.1.2, the ansatz unitary is of the form $U_{\mathrm{UCC}}(\boldsymbol{\theta}) = \exp[\hat{T}(\boldsymbol{\theta}) - \hat{T}^{\dagger}(\boldsymbol{\theta})]$, which can be implemented by decomposing it via the Trotterization procedure (see 6.2.3). Note, that $U_{\mathrm{UCC}}(\boldsymbol{\theta})$ is equivalent to an evolution of constant time, in contrast with the exponentially long time evolutions required for QPE. Therefore, the length of the resulting Trotterization circuit is greatly reduced here. The reference state $|\psi_{ref}\rangle$ is typically very simple to prepare (e.g. is a state in the computational basis), so that U_{ref} is usually given by simple local rotations of the qubits or the identity [29].

Once the trial state $|\psi(\boldsymbol{\theta})\rangle$ is prepared, evaluating the expectation value of the Hamiltonian $E(\boldsymbol{\theta}) = \langle\psi(\boldsymbol{\theta})|\,\hat{H}\,|\psi(\boldsymbol{\theta})\rangle$ in an efficient fashion is not straightforward. In order to do it, we restrict our attention to Hamiltonians that allow an efficient expectation value estimation on quantum hardware [17, 30]. That is, Hamiltonians that can be decomposed into a set of Hermitian operators

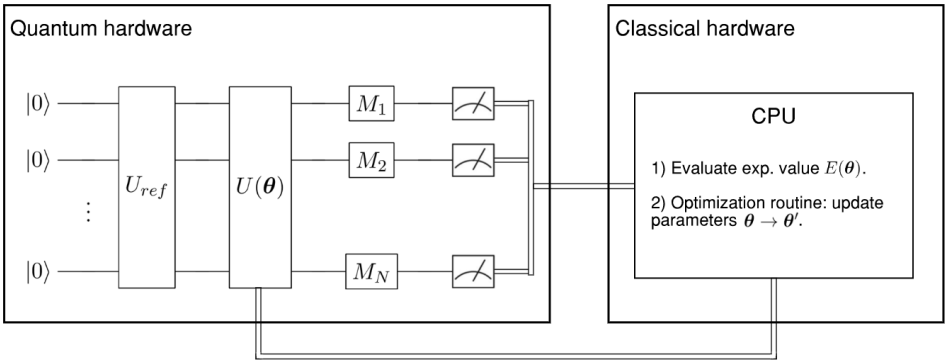


Figure 6.19: Schematic of the VQE algorithm. The quantum hardware is used to prepare the trial states $|\psi(\boldsymbol{\theta})\rangle$ according to the ansatz, and perform local rotations and measurements to evaluate the expectation values of Pauli operators. The individual measured expectation values are fed into a classical processor which computes the expectation value of the Hamiltonian $E(\bar{\boldsymbol{\theta}})$. This objective function value is then fed into an optimisation routine, which produces an updated set of parameters $\boldsymbol{\theta}'$ as input for the quantum circuit in the next step of the optimisation routine. The procedure is repeated until the energy converges to a minimum.

$$\hat{H} = \sum_{i=1}^{K} h_i \hat{O}_i \tag{6.5.3}$$

where the number of operators K is only polynomial with the size of the system M, and each Hermitian operator $\hat{O}_i$ can be efficiently measured on the quantum hardware (e.g. contains only local terms). If such a decomposition is possible then also the expectation value of the Hamiltonian $\langle \hat{H} \rangle$ can be evaluated efficiently by writing it as a sum of a polynomial number of "easy"terms

$$\langle \hat{H} \rangle = \sum_{i=1}^{K} h_i \langle \hat{O}_i \rangle. \tag{6.5.4}$$

Hermitian operators that can be efficiently measured are typically products of local operators. The decomposition is thus usually done in the basis of Pauli operators[9]

$$\hat{H} = \sum_{i,\alpha} h_i^{\alpha} \hat{\sigma}_i^{\alpha} + \sum_{i,j,\alpha,\beta} h_{ij}^{\alpha,\beta} \hat{\sigma}_i^{\alpha} \otimes \hat{\sigma}_j^{\beta} + \ldots, \tag{6.5.5}$$

9 As products of Pauli operators form a basis in the operator space, it is always possible to make a decomposition as in eq. (6.5.5). However, in general such a decomposition involves an exponential number of terms, rather than polynomial. Intuitively, a polynomial decomposition can be obtained in most systems of practical interest by exploiting localities in the Hamiltonian [10, 30].

where $\alpha, \beta, \ldots \in \{x, y, z\}$ index different Pauli matrices. The expectation value of the energy $E(\boldsymbol{\theta})$ is therefore estimated taking measurements of individual tensor products of Pauli terms appearing in eq. (6.5.5), which only require local rotations and measurements of the qubits (see Fig. 6.19). The results of the indivual terms are then fed to a classical processor which computes the weighted sum in eq. (6.5.5), obtaining $E(\boldsymbol{\theta})$.

The expectation value $E(\boldsymbol{\theta})$ is used as objective function in a minimisation algorithm run on the classical processor. Different types of optimisation algorithms can be used, e.g. Nelder-Mead or particle swarm approaches, with the optimal choice likely to depend on the particular problem targeted and the noise affecting the experiment [4, 30]. The updated trial parameter values $\boldsymbol{\theta}'$ obtained from the classical optimisation algorithm are fed back into the quantum circuit which restarts the routine. This procedure is repeated until the energy converges to a minimum, which represents our approximation of the ground state energy.

6.5.2 *Advantages and open questions for VQE*

The main advantages of VQE come from the fact that no time-evolution is required for the trial state, which is instead measured in a local basis as soon as it is prepared, which takes $\mathcal{O}(1)$ time and immediately destroys all coherence. This contrasts with PEA-based methods, where, after the trial state is prepared, a long series of gates has to be applied to perform the controlled-U^{2^k} operations while maintaining the coherence of the state. Avoiding the need of long coherence times for the qubits can greatly enhance the practicability of the approach on current devices affected by this type of noise, as was recently experimentally shown in a direct comparison between the VQE and PEA-based approaches for a small-scale molecule [27]. Moreover, VQE allows some flexibility for choosing an ansatz which suits the particular experimental platform used. This feature can add additional robustness to other types of noise, such as gate errors [27, 30]. Even in cases where some form of error-correction needs to be implemented, as is expected to be required for future large-scale implementations, this robustness could translate to significantly reduced overheads compared to PEA-based methods. How the noise-robustness advantage compares in practice to new approaches to phase estimation, such as the RFPE algorithm discussed in section 6.3, is not clear though.

However, some drawbacks are also introduced, compared to PEA-based methods. First, the estimation of the energy in VQE is performed via repeated experiments, which requires a number of resources that scales as $\mathcal{O}(1/\epsilon^2)$ (shot-noise scaling) with the precision ϵ in the estimation, compared to $\mathcal{O}(1/\epsilon)$ in PEA-based methods (Heisenberg scaling) [17]. A second, more crucial, issue is that while in PEA the collapse of the wavefunction to an exact eigenstate is achieved even with minimal requirements for the ansatz (i.e. non-zero overlap with the target state), in VQE the role of the ansatz is

central in obtaining a precise approximation of the energy. In other words, while PEA is equivalent to using a full configuration interaction description (see section 6.1.2), in VQE it is not known in general if the chosen ansatz will allow a correct representation of the target eigenstate of the system [17, 30]. Although it is widely believed that the class of ansätze that can be efficiently encoded on a quantum processor can explore a wider space with respect to what is possible on classical machines, the development of such ansätze currently remains an open problem [4, 30–32]. In fact, while an efficient eigenstate preparation is very unlikely to be possible for arbitrary Hamiltonians[10], the hope is that such polynomial ansätze do exist for quantum chemistry problems of practical interest [4, 30–32].

Another issue for VQE is to determine the scaling of the number of iterations of the variational minimisation algorithm required to converge to the optimal solution. This quantity is very difficult to study analytically, as it depends on a large number of factors, such as the properties and degeneracies of the system Hamiltonian, the ansatz chosen, and the classical algorithm used for the minimisation. The hope is that, given that the dimension of the ansatz parameters space grows only polynomially with the number of orbitals, also the number of iterations required is polynomial for practical cases. However, given that the energy landscape in the ansatz parameter is highly non-convex, instances may arise where the algorithm gets stuck in local minima, preventing a correct convergence in polynomial time. Recent experimental results have provided some evidence that an efficient convergence of the minimisation algorithm is achievable implementing VQE for small molecules [30–32], which somehow suggest that it can be the case also for more complex systems.

The open questions described above makes it hard to asses the asymptotic behaviour of the VQE approach, and its scalability for the moment has to rely on a number of controversial assumptions. It seems likely that only by testing the algorithm on molecules of increasing size with actual large-scale quantum processors we will be able to assess these questions.

6.6 Witness-assisted variational eigensolvers for excited states

Original approaches to VQE, where states are found by minimising the energy as in eq. (6.5.1), mainly focused on finding ground-states of the systems. Generalising the approach to eigenstates with higher energies (i.e. excited states) while maintaining the practicality of VQE, is quite an interesting problem.

10 One of the most famous Hamiltonians for which it was shown that no efficient eigenstate preparation is possible, under well-established computational assumptions (i.e. assuming $\mathsf{NP} \not\subseteq \mathsf{BQP}$) is the Ising model Hamiltonian [33].

Excited states have widespread applicability in quantum chemistry. For example, characterising the eigenspectrum of molecular Hamiltonians can be used to understand reaction rates in quantum chemistry, and in chapter 5 we saw applications for molecular optical spectra. In particular, characterisation of excited states is required to study energy and charge transfer processes such as those in bulk heterojunction solar cells or photosynthetic light harvesting complexes [34–36]. As for ground states, in general finding excited states is also complex on classical machines. In fact, excited states are often hard to access also in instances where classical methods can successfully describe the ground state (e.g. for weakly interacting Hamiltonians) [35, 37], increasing the interest towards developing quantum methods for their efficient search. Clearly, as the size of the Hamiltonian, and consequently the elements in its spectrum, increases exponentially with the size of the system, finding all eigenstates becomes rapidly computationally intractable also for quantum approaches. However, the excited states of practical interest are typically those closer in energy to the ground state [35]. The ability to efficiently access a small (possibly constant) number of low energy excited stated would thus be sufficient for a wide range of applications.

A possible approach to variationally target excited states through VQE, inspired by classical methods [38], is the folded spectrum method [17, 30]. The idea is to perform the variational minimisation in eq. (6.5.1) using a spectrally transformed Hamiltonian $\hat{H}' = (\hat{H} - \gamma\mathbb{1})^2$, with γ a real parameter. However, as the effective Hamiltonian H' contains the squared original Hamiltonian H^2, the number of terms in it (see eq. (6.5.5)) grows quadratically. As the expectation values of these terms need to be estimated one by one in VQE, the computational cost is increased quadratically when adopting the folded spectrum approach. This issue significantly increases the complexity of implementing the procedure even for small-scale implementations, which, despite their practical importance, prevented the investigation of excited states in the first experimental VQE demonstrations [17, 27].

In this section we will describe an alternative approach to VQE which we developed explicitly to target excited states, called Witness-Assisted Variational EigenSolver (WAVES). We will then present a small-scale experimental photonic implementation of the WAVES algorithm on the integrated photonic device presented in section 6.4. Thanks to the enhanced practicability of WAVES, in the experiment we targeted both the ground and the first excited state of a simple system. This represented the first experiment where an excited state was accessed on a digital quantum processor[11].

11 Following the demonstration reported here, another experiment was reported a few months later where excited states were variationally simulated using the folded spectrum approach in a superconducting qubit architecture [39].

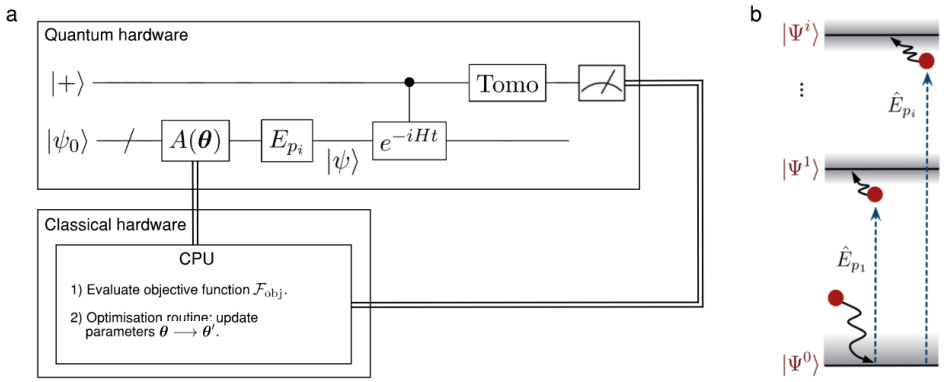


Figure 6.20: The WAVES protocol. **a,** Schematic of the quantum circuit for WAVES, which allows to efficiently measure the objective function $\mathcal{F}_{\rm obj} = \mathcal{E} + T\mathcal{S}$ from the tomography of the single control qubit. Initially, a variational search for the ground state $|\Psi_0\rangle$ of the Hamiltonian is performed by preparing trial states via the ansatz $A(\boldsymbol{\theta})$ with no perturbation (i.e. $E_{p_0} = \mathbb{1}$), and driving the optimisation algorithm via a classical processor minimising $\mathcal{F}_{\rm obj}$. The ground state is then used as a reference for the excited states search, where an initial guess for an excited state $|\Psi_i\rangle$ is given by a perturbation E_{p_i} on the ground state and then refined using the same circuit. The variational search of the excited state in now performed by exploiting the eigenstate witness $\mathcal{F}_{\rm obj} = \mathcal{S}$ (high-T limit of the objective function). **b,** Diagram representing schematically the intuition behind the proposed approach, where initial guesses of excited states are variationally refined using the eigenstate witness.

6.6.1 *The WAVES algorithm*

The WAVES algorithm is based on the introduction of the concept of an *eigenstate witness*, that is a quantity which identifies whether a test state is close to being an eigenstate or not. While no efficient analogues are known for classical algorithms, we will see how such a quantity can be defined and efficiently measured on quantum hardware. In particular, this witness will be based on the entropy acquired by a quantum register whose time evolution is controlled by an ancillary single qubit. A schematic of the hybrid quantum-classical circuit for WAVES is shown in Fig. 6.20. The search for a target eigenstate proceeds by preparing trial states $|\psi\rangle_S$ in the system register of n qubits, according to the ansatz, and initialising the single control qubit in the $|+\rangle_C = (|0\rangle + |1\rangle)/\sqrt{2}$ state. The combined state $|+\rangle_C \otimes |\psi\rangle_S$ then undergoes a controlled unitary operation, where the unitary $U = \exp(-i\hat{H}t)$ represents a time evolution of the trial state according to the system Hamiltonian $\hat{H}$ for a pre-fixed amount of time t, yielding the state

$$\frac{1}{2}(|0\rangle_C |\psi\rangle_S + |1\rangle_C e^{-i\hat{H}t} |\psi\rangle_S) \tag{6.6.1}$$

After the unitary, the individual state of the emerging control qubit is given by the partial trace over the system registry $\rho_C = \mathrm{Tr}_S(\rho)$, which can be explicitly written as

$$\rho_C = \frac{1}{2}\begin{pmatrix} 1 & (\langle\psi| e^{-iHt} |\psi\rangle)^* \\ \langle\psi| e^{-iHt} |\psi\rangle & 1 \end{pmatrix}. \tag{6.6.2}$$

The reduced density matrix ρ_C is reconstructed via state tomography of the control qubit, which requires a small number (independent of the size of the system) of measurements, and can be used to extract two important quantities.

The first quantity of interest is the Von Neumann entropy of the reduced state $\mathcal{S}(\rho_C)$, which is defined as $\mathcal{S}(\rho) = -\mathrm{Tr}(\rho \ln \rho)$. As the Von Neumann entropy has the property of being the same for each reduced subsystem of a bipartite state (see e.g. [10]), it holds that $\mathcal{S}(\rho_C) = \mathcal{S}(\rho_S)$, meaning that the actual entropy of the system register $\mathcal{S}(\rho_S)$ can be measured via reconstructing the single qubit tomography ρ_C. The entropy $\mathcal{S}(\rho_S)$ plays a crucial role in the algorithm as it acts as an eigenstate witness: it is zero if the trial state is an eigenstate of the Hamiltonian. The idea behind this is quite simple. If the trial state $|\psi\rangle_S$ is an eigenstate of the Hamiltonian, then the emerging state in eq. (6.6.1) will be the separable state $(|0\rangle + e^{-i\lambda t} |1\rangle) \otimes |\psi\rangle_S$, with λ the associated eigenvalue. Being a separable state, the partial state of each subsystem is a pure state, i.e. with zero Von Neumann entropy ($\mathcal{S}(\rho_C) = \mathcal{S}(\rho_S) = 0$). On the contrary, if $|\psi\rangle_S$ is not an eigenstate, some entanglement will be present in the output state, meaning non-zero entropy $\mathcal{S}(\rho_S) > 0$ for the reduced state of the system. With such a property, this eigenstate witness allows us to target eigenstates variationally, both ground and excited states, by minimising $\mathcal{S}(\rho_S)$.

A second important quantity that can be estimated from the reconstructed single qubit matrix ρ_C is an energy estimator

$$\mathcal{E} = -\arg(\langle\psi| e^{-i\hat{H}t} |\psi\rangle)/t \tag{6.6.3}$$

which can be evaluated from the off-diagonal elements of the reconstructed ρ_C in eq. (6.6.2).

To variationally find the ground state we can then minimise the energy estimate $\mathcal{E}$, as for standard VQE. However, now we can also simultaneously minimise the entropy $\mathcal{S}(\rho_S)$. This allows us to computationally interpret the task as a minimisation of a physically motivated objective function $\mathcal{F}_{\mathrm{obj}} = \mathcal{E} + T\mathcal{S}(\rho_S)$, where T is a parameter that trades off between energy optimisation and entropy optimisation. Note that interestingly $\mathcal{F}_{\mathrm{obj}}$ does have a physical analogy to the *Helmotz free energy*, the minimisation of which provides equilibrium states in statistical mechanics (see e.g. [40]). However, given a sign difference between these free-energies, in this analogy T plays the somewhat contrived

role of a negative temperature. In the high T limit, where the entropy becomes dominant, the objective function $\mathcal{F}_{\text{obj}}$ also permits one to identify excited states. However, as the entropy minimisation procedure is more likely to converge to the eigenstate which has the largest overlap with the initial trial state (a situation similar to the PEA), when looking for a particular excited state we need to start from an approximate guess. For this reason, defining an initial reference state $|\psi_0\rangle$ (usually obtained by Hartree-Fock appoximation) and the vector $\boldsymbol{\theta}$ as the list of parameters describing the ansatz-based state preparation $A(\boldsymbol{\theta})$, i.e. $|\psi\rangle_S = A(\boldsymbol{\theta})\,|\psi_0\rangle$, the WAVES protocol proceeds in two phases:

1. Ansatz-based variational search for the ground state, obtaining the parameters $\boldsymbol{\theta}_g$ that minimize the objective function $\mathcal{F}_{\text{obj}}$ and the associated unitary for the ground state $A_g = A(\boldsymbol{\theta}_g)$.

2. Witness-assisted variational search for excited states, starting with initial guesses obtained from the ground state reference. The unitary for the initial approximate guess of the i-th target excited state is obtained via $E_{p_i}A(\boldsymbol{\theta}_g)$, with E_{p_i} a system dependent perturbation. Using as initial starting point of the ansatz parameters $\boldsymbol{\theta} = \boldsymbol{\theta}_g$, the search for the excited state is done by variationally finding the parameters $\boldsymbol{\theta}_{e_i}$ that minimise $\mathcal{F}_{\text{obj}}$ in the high T limit (where the entropy is dominant), obtaining the unitary for the target excited state $A_{e_i} = E_{p_i}A(\boldsymbol{\theta}_{e_i})$.

An important point is now how to choose the excitation operators E_{p_i} used in the variational search for excited states. Different approaches to do this have been recently proposed in Ref. [41]. A first approach is to work in the Hartree-Fock approximation. In this case, considering a general many-body Hamiltonian $H = H_0 + V$, with $H_0 = \sum_i \xi_i a_i^\dagger a_i$ the single-particle contributions and V the interaction terms, we consider H_0 to be dominant. In such a scenario, a transition from the ground state to an excited state can be approximated by the action of a sequence of single excitation operators $a_i^\dagger a_j$, with a corresponding unitary $E_p = \exp[\pi(a_i^\dagger a_j - a_j^\dagger a_i)/2]$. This approximation for the perturbations is believed to be sufficient to have an adequately good initial guess in most practical cases. On the other hand, if hard instances arise where it is not sufficient, it can be refined with more sophisticated methods, such as multi-configuration self-consistent-field approximations [4, 32, 41].

6.6.2 *Performance of the algorithm*

To summarise, while the initial ansatz-based trial state preparation is common between WAVES and standard VQE, the differences are in the requirement of the controlled

unitary operation with an ancillary qubit, the measurement performed on the single qubit instead of the system register, and the objective function used in the variational search.

The controlled $\exp(-iHt)$ unitary requirement implies that, in contrast to VQE, additional operations need to be done after the state preparation. Therefore, longer coherence times are required for WAVES compared to VQE. However, in contrast to the unitaries performed in PEA-based methods where t rapidly explodes, the evolution time t is fixed and small. In particular, in order to avoid degeneracies in the energy estimation, we require $t \in (0, \pi/2\|H\|)$, with $\|H\|$ the Frobenius norm of the Hamiltonian [2]. Therefore, the requirement of extremely long coherence times is avoided in WAVES, as well as in standard VQE. Roughly speaking, the circuit depth required to implement the Trotter-Suzuki decomposition of the unitary has approximately the same size of the Trotterised implementation of the ansatz unitary in, for example, the unitary coupled cluster ansatz case[12]. The circuit depth for WAVES should thus be a small constant factor (dependent on the particular implementation) larger than for VQE, maintaining a significant advantage compared to PEA-based methods.

The measurement part for WAVES is quite simple, requiring only detection of the single control qubit. The objective function can be inferred from the reconstructed single qubit tomography of the control qubit, which thus requires to repeat the experiment $\mathcal{O}(1/\epsilon^2)$ to obtain the objective function up to a precision ϵ, which is the same scaling as standard VQE. Crucially, this scaling is independent on whether the ground or excited states are targeted, avoiding the quadratic overheads required for the folded spectrum method in standard VQE [30]. This represents the main advantage of the WAVES approach with respect to previous methods.

In terms of total computational complexity, WAVES is currently affected by the same limitations and open questions as for the VQE approaches. Namely, its scalability needs to assume that polynomial ansätze and efficient convergence of the variational optimisation algorithm are possible for cases of practical interest. Under these assumptions, both VQE and WAVES are scalable approaches. The use of an objective function based not only on the energy but also on the entropy could provide some advantages for WAVES over VQE in terms of avoiding local minima in the parameter space. While preliminary tests on small systems seem to suggest that the parameter T can indeed be engineered to do that, we still don't know if that's also the case for molecules with, say, tens of orbitals. Investigating this aspect could be an interesting further research direction.

12 The number of operations required to perform the Trotter-Suzuki decomposition for both the ansatz unitary and the controlled unitary is roughly given by $n^{5.5}$ for quantum chemistry problems, as investigated in [42]. Here n is the number of orbitals of the simulated molecule.

6.7 Experimental implementation of WAVES

A proof-of-principle experimental demonstration of WAVES was performed on the two-qubit silicon quantum photonic processor that was already presented in section 6.4, where we described its capability to perform arbitrary state preparation on the system qubit, arbitrary controlled unitary operations between two qubits, and arbitrary measurements on the control qubit (see Fig. 6.10). These are exactly the operations required to implement the WAVES circuit in Fig 6.20a for a two-dimensional (single-qubit) Hamiltonian. Note that, similarly as for the RFPE implementation in section 6.4, here the state preparation and the controlled operation are implemented by direct exponentiation of the ansatz and Hamiltonian operators, mapping them directly into the chip phases. While this is possible in our small-scale experiment, scalable implementations on larger devices will need to adopt the Trotter-Suzuki decomposition to efficiently implement such operations.

6.7.1 *Hamiltonian system investigated*

In our implementation, we used WAVES to calculate the eigenspectrum of a simplified exciton transfer Hamiltonian of two Chlorophyll units in the 18-mer ring of the LHII complex [43]. In the basis of localised excitons on each Chlorophyll unit (see insets of Fig. 6.21a and Fig. 6.22a), the exciton transfer Hamiltonian is given by the 2×2 matrix

$$\hat{H} = \begin{pmatrix} \alpha & \beta \\ \beta & \alpha \end{pmatrix}, \tag{6.7.1}$$

where $\alpha = 1.46$ eV is the energy of the exciton on one of the Chlorophyll units and $\beta = 0.037$ eV is the interaction between the excitons arising from the transition dipole between the two units. The Hamiltonian in the qubit basis, obtained via compact mapping [30], is encoded in the single qubit operator $\hat{H}_{\text{qubit}} = \alpha\hat{\mathbb{1}} + \beta\hat{\sigma}_x$, with $\hat{\sigma}_x$ the Pauli x operator. This Hamiltonian has ground (excited) state $|-\rangle$ ($|+\rangle$), with associated energy $\alpha - \beta$ ($\alpha + \beta$). For this system, the perturbation unitary for the excited state corresponds to $\hat{E}_p = \exp(i\pi\hat{\sigma}_z/2)$. Given the small size of the problem, we used a full-configuration interaction ansatz for the state preparation (see section 6.1). In order to better resolve the energy difference between the two eigenstates $2\beta \ll \alpha$, we rescale the Hamiltonian via the transformation $\hat{H} \to \hat{H} - \gamma\hat{\mathbb{1}}$, which maintains the same eigenstates as $\hat{H}$ but rescales its eigenvalues λ_i as $\lambda_i \to \lambda_i - \gamma$. In other words, the parameter γ simply redefines arbitrarily the energy zero. To mimic a realistic problem where rather poor guesses for the exact eigenvalues are possible, in the particular implementation reported here a rather arbitrary value of $\gamma = 1.27$ eV was used.

As described above, due to the fact that the energy estimation in WAVES adopts the form of eq.(6.6.3), a restriction on the value of t is needed to avoid issues due to the

2π periodicity of the $\arg(\cdot)$ function. This is a restriction that we already encountered for the PEA-based methods, and can be simply dealt with by choosing t small enough. In particular, we need $t \in (0, \pi/2\|H\|)$, which still requires very rough preliminary estimates for the Hamiltonian eigenvalues (e.g. those classically available via mean-field methods) [2]. In the same spirit as for the choice of γ, we picked a rather arbitrary value $t = 26\ \mathrm{eV}^{-1}$, which satisfies the above requirements for the Hamiltonian under investigation.

6.7.2 *Optimisation algorithm*

A wide range of algorithms can be used for the classical optimisation procedure to minimise the objective function $\mathcal{F}_{\mathrm{obj}}$, which can be mostly classified as gradient-based and gradient-free methods [4, 30, 32]. Here, a gradient–free particle swarm optimisation algorithm was used, inspired by the Bayesian inference techniques adopted in RFPE (see section 6.3), thus retaining part of their noise robustness. The idea is to have a swarm of trial states (particles) randomly sampled from a prior distribution. The objective function $\mathcal{F}_{\mathrm{obj}}$ is measured for each particle, and the result is used to update the probability distribution. A new swarm is drawn from the updated distribution to perform further steps, until a convergence criterion is triggered. The last inferred posterior provides the estimate for the optimal state and the corresponding objective function. In the case of a ground-state search, we start sampling a swarm Ξ composed of N particles (i.e. a set of N parameters values $\{\boldsymbol{\theta}_j\}$) from the parameter space according to a uniform distribution (we assume no prior knowledge of the ground-state). For each of the particles, the objective function $\mathcal{F}_{\mathrm{obj}}$ is efficiently obtained from the quantum device. Once the objective function has been obtained for all the particles, the algorithm rejects the subset of size $N - N_\times$ of them[13] with associated higher values of $\mathcal{F}_{\mathrm{obj}}$, reducing the set to Ξ'. This smaller set of particles is used to classically compute a weighted[14] mean $\mu_{\theta'}$ and covariance $\Sigma_{\theta'}$ of a Gaussian distribution on the parameters space. The set of particles is then repopulated by adding back $N - N_\times$ new particles drawn from the (last updated) Gaussian distribution $\mathcal{N}(\mu_{\theta'}, \Sigma_{\theta'})$, forming the new set of N particles Ξ to be used in the next iteration of the algorithm.

These iterations are repeated until convergence is achieved, which happens either when the Gaussian distribution covariance norm is sufficiently peaked (i.e. the spread of the distribution is lower than a pre-defined accuracy), or when fluctuations of $\mathcal{F}_{\mathrm{obj}}$ in the last few steps are smaller than a pre-determined threshold (indicating that the algorithm

13 In the experiment we used $N_\times = \lceil \sqrt{N} \rceil$, but for more challenging implementations it is likely that $N_\times \in \mathcal{O}(N)$ will be required.

14 The weights used to calculate the mean and covariance of the posterior distribution are chosen to be proportional to the objective function $\mathcal{F}_{\mathrm{obj}}$ associated to each particle.

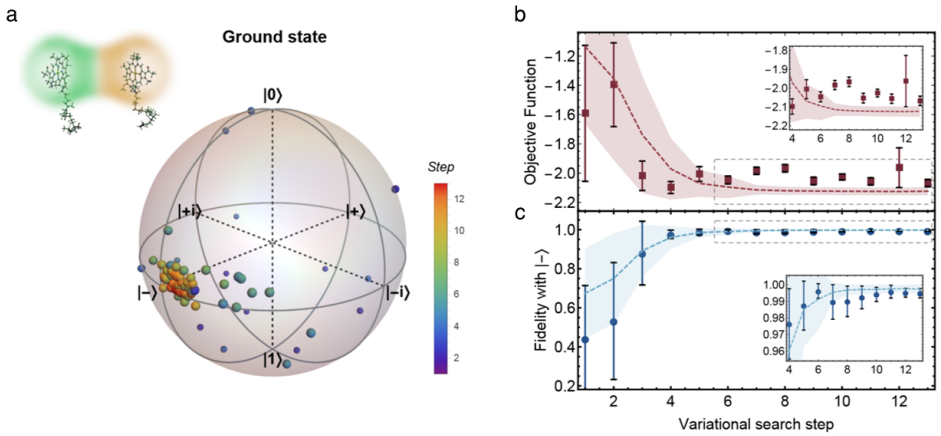


Figure 6.21: Experimental variational search on the ground state via WAVES. **a,** Colour-coded evolution of the particle swarm for the WAVES search of the ground state ($|-\rangle$) shown on the Bloch sphere. Different colours correspond to different steps of the search protocol. Inset: Graphic representation of the ground-state (anti-symmetric) for a single-exciton transfer Hamiltonian between two Chlorophyll sites. Panels **b** and **c** report the evolution of the objective function $\mathcal{F}_{\text{obj}}$ and of the fidelity ($F = |\langle\psi|\psi_{\text{ideal}}\rangle|^2$), respectively, as a function of the search step. Error bars are given by the variance of the particles distribution and photon Poissonian noise. Dashed lines are numerical simulations of the performance of the algorithm, averaged over 1000 runs, with shaded areas representing a 67.5% confidence interval. Insets: behaviour close to convergence.

is drawing new particles in a plateau of the objective function).

The swarm optimisation in the case of excited-state searches is fully equivalent, with the only difference that the initial distribution of the swarm Ξ is assumed as a Gaussian distribution centred around the initial guess provided by the appropriate excitation operator.

6.7.3 *Results of the experiment*

The experimental results of the WAVES protocol for the ground and excited states variational search are shown in Fig. 6.21 and Fig. 6.22, respectively. The search is performed by optimising the objective function $\mathcal{F}_{\text{obj}} = \mathcal{E} + T\mathcal{S}(\rho_S)$, where for the parameter T a small value of $T = 1.25$ was used for the ground state search, while for the excited state case we used $\mathcal{F}_{\text{obj}} = \mathcal{S}(\rho_S)$, i.e. adopting a large-$T$ limit. The minimisation of the objective function was performed in both cases adopting the particle-swarm method outlined above with $N = 8$ particles. In Fig. 6.21a and Fig. 6.22a the colour-coded evolution of the swarm is reported, achieving rapid convergence of the particle distribution towards the expected eigenstates of the Hamiltonian: the ground state $|-\rangle$ and the first

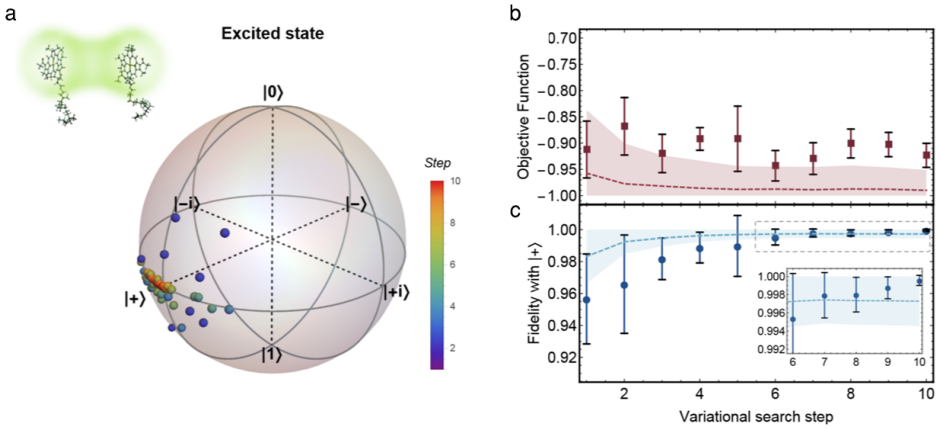


Figure 6.22: Experimental variational search on the excited state via WAVES. **a,** Colour-coded evolution of the particle swarm for the WAVES search of the excited state ($|+\rangle$) shown on the Bloch sphere. Different colours correspond to different steps of the search protocol. Inset: Graphic representation of the excited-state (symmetric) for a single-exciton transfer Hamiltonian between two Chlorophyll sites. Panels **b** and **c** report the evolution of the objective function $\mathcal{F}_{\text{obj}}$ and of the fidelity ($F = |\langle\psi|\psi_{\text{ideal}}|^2$), respectively, as a function of the search step. Error bars are given by the variance of the particles distribution and photon Poissonian noise. Dashed lines are numerical simulations of the performance of the algorithm, averaged over 1000 runs, with shaded areas representing a 67.5% confidence interval. Insets: behaviour close to convergence.

excited state $|+\rangle$. For the ground state search, pessimistically assuming no pre-existing knowledge of the system, the prior is initialised to span uniformly the sub-section of the Hilbert space identified by the ansatz. For the excited state, instead, the search is initialised with the guessed state obtained by applying $\hat{E}_p$ to the ground state. As shown in Fig. 6.21b and Fig. 6.22b, within 10-13 search steps the $\mathcal{F}_{\text{obj}}$ is converges to its minimal value, which correctly corresponds to the ground state and excited state, respectively. In fact, Fig. 6.21c and Fig. 6.22c report that the eigenstate estimator, provided by the mean of the particles distribution, achieves a high overlap with the eigenstate targeted: fidelities of $99.48 \pm 0.28\%$ with the ground state and $99.95 \pm 0.05\%$ with the excited one are achieved. All uncertainties are given by the variance of the prior distributions: a well motivated error bar is among the amenable features deriving from the adoption of a swarm optimisation method. Imperfect measurements of $\mathcal{F}_{\text{obj}}$, more evident in the regime close to convergence (see insets of Fig. 6.21c and Fig. 6.21c), are due mainly to experimental noise in the phase settings, given by the residual thermal cross-talk among the phase-shifters already discussed in section 6.4. The fast convergence of the algorithm, however, indicates a good robustness of the protocol to this kind of experimental noise.

6.7.4 *Experiment discussion*

A proof-of-principle optical implementation of the WAVES algorithm on a silicon quantum photonic chip was demonstrated for a simplified exciton transfer Hamiltonian, obtaining its eigenstates with high fidelities. Together with the description of the eigenstates, as parametrised by the ansatz, WAVES can also return the corresponding eigenvalue estimation via $\mathcal{E}$. However, as for standard VQE, the precision ϵ for such eigenvalue estimation requires a number of measurements $\mathcal{O}(1/\epsilon^2)$, and is highly dependent on the reliability of the ansatz to correctly describe the state. Alternatively, one can use the variational search in WAVES as a state preparation stage and then feed the obtained state to a IPEA or RFPE circuit to extract the eigenphase. In this way, the shortcomings present in the variational ansatz are avoided and a quadratic resource scaling enhancement $\mathcal{O}(1/\epsilon)$ is achieved for the measurement. This, however, would again come at the expense of long coherence times required for the PEA implementation.

The experimental demonstration involved a very simple two-dimensional Hamiltonian, for which WAVES achieved very promising results. However, to investigate the actual scalability of the protocol, tests on larger systems will be important. Preliminary tests, reported in [2], demonstrate that WAVES performs well for quantum chemical problems involving up to eight-qubit Hamiltonians, with the variational search efficiently finding target eigenstates (for both ground and excited states) with an average fidelity of $\simeq 99\%$. The development of complex quantum devices will be ultimately required to test WAVES for systems at a scale where new challenging problems may arise.

6.8 Discussion

In this chapter we have discussed different types of digital quantum algorithms for the simulation of quantum chemistry, and the small-scale implementation of two of these algorithms, RFPE and WAVES, on a silicon photonic chip. Current algorithms for digital quantum simulation can be divided in two main areas: PEA-based algorithms and variational algorithms. The first group, PEA-based algorithms, have the important advantage of having loose requirements on the state preparation and deterministically provide the energy estimate, which strongly suggest their scalability for practical quantum chemistry problems. However, although Bayesian approaches significantly improve their noise robustness, they need deep quantum circuits, which do not seem to be suitable for near-term quantum hardware. On the other hand, variational approaches need much shorter circuit depth, but their scalability strongly relies on the reliability of the ansatz and the optimisation procedure, which still remain important open problems.

Acknowledgements

The author's contributions to the material reported in this chapter are the experimental implementation and data analysis for the RFPE and WAVES algorithms. The author also provided inputs to the development of the WAVES algorithm, and coded a first version of the swarm optimisation algorithm employed. The experiments and the analysis were done in cooperation with Dr. Raffaele Santagati, Dr. Jianwei Wang and Mr. Antonio A. Gentile, joined by Dr. Xiao-Xi Zhou, Dr. Nathan Wiebe and Dr. Jarrod R. McClean for the development of the WAVES protocol. The chip used for the experiments was fabricated by Toshiba. In all implementations the mapping of chemical systems to qubits was done by Dr. David P. Tew. Dr. Nathan Wiebe additionally provided theoretical support for the RFPE experiment, for which we also acknowledge useful discussions with Dr. Chris Granade. The experiment regarding the Bayesian approach to phase estimation is published in Physical Review Letters [1], while the theory of the WAVES algorithm and its experimental implementation are published in Science Advances [2].

References

1. Paesani, S., Gentile, A. A., Santagati, R., Wang, J., Wiebe, N., Tew, D. P., O'Brien, J. L. & Thompson, M. G. Experimental Bayesian Quantum Phase Estimation on a Silicon Photonic Chip. *Phys. Rev. Lett.* **118,** 100503. https://link.aps.org/doi/10.1103/PhysRevLett.118.100503 (2017).
2. Santagati, R., Wang, J., Gentile, A. A., Paesani, S., Wiebe, N., McClean, J. R., Morley-Short, S., Shadbolt, P. J., Bonneau, D., Silverstone, J. W., *et al.* Witnessing eigenstates for quantum simulation of Hamiltonian spectra. *Science Advances* **4,** eaap9646 (2018).
3. Born, M. & Oppenheimer, R. Zur quantentheorie der molekeln. *Annalen der Physik* **389,** 457–484 (1927).
4. McArdle, S., Endo, S., Aspuru-Guzik, A., Benjamin, S. & Yuan, X. Quantum computational chemistry. *arXiv:1808.10402* (2018).
5. Helgaker, T., Jorgensen, P. & Olsen, J. *Molecular electronic-structure theory* (John Wiley & Sons, 2014).
6. Romero, J., Babbush, R., McClean, J. R., Hempel, C., Love, P. J. & Aspuru-Guzik, A. Strategies for quantum computing molecular energies using the unitary coupled cluster ansatz. *Quantum Science and Technology* **4,** 014008 (2018).
7. Jordan, P & Wigner, E. Über das Paulische Äquivalenzverbot. *Zeitschrift für Physik* **47,** 631–651 (1928).
8. Aspuru-Guzik, A., Dutoi, A. D., Love, P. J. & Head-Gordon, M. Simulated quantum computation of molecular energies. *Science* **309,** 1704–1707 (2005).

9. Shor, P. Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer. *SIAM J. Comput.* **26,** 1484 (1997).

10. Nielsen, M. A. & Chuang, I. *Quantum computation and quantum information* (Cambridge University Press, 2002).

11. Dobšíček, M., Johansson, G., Shumeiko, V. & Wendin, G. Arbitrary accuracy iterative quantum phase estimation algorithm using a single ancillary qubit: A two-qubit benchmark. *Physical Review A* **76,** 030306 (2007).

12. Kitaev, A. Y. Quantum measurements and the Abelian stabilizer problem. *arXiv:quant-ph/9511026* (1995).

13. Lanyon, B. P., Whitfield, J. D., Gillett, G. G., Goggin, M. E., Almeida, M. P., Kassal, I., Biamonte, J. D., Mohseni, M., Powell, B. J., Barbieri, M., *et al.* Towards quantum chemistry on a quantum computer. *Nature chemistry* **2,** 106 (2010).

14. Lloyd, S. Universal quantum simulators. *Science,* 1073–1078 (1996).

15. Wecker, D., Bauer, B., Clark, B. K., Hastings, M. B. & Troyer, M. Gate-count estimates for performing quantum chemistry on small quantum computers. *Phys. Rev. A* **90,** 022305. https://link.aps.org/doi/10.1103/PhysRevA.90.022305 (2014).

16. Xu, J.-S., Yung, M.-H., Xu, X.-Y., Boixo, S., Zhou, Z.-W., Li, C.-F., Aspuru-Guzik, A. & Guo, G.-C. Demon-like algorithmic quantum cooling and its realization with quantum optics. *Nature Photonics* **8,** 113 (2014).

17. Peruzzo, A., McClean, J., Shadbolt, P., Yung, M.-H., Zhou, X.-Q., Love, P. J., Aspuru-Guzik, A. & O'Brien, J. L. A variational eigenvalue solver on a photonic quantum processor. *Nature communications* **5,** 4213 (2014).

18. Mitzenmacher, M. & Upfal, E. *Probability and computing: Randomized algorithms and probabilistic analysis* (Cambridge university press, 2005).

19. Wiebe, N. & Granade, C. Efficient Bayesian Phase Estimation. *Phys. Rev. Lett.* **117,** 010503. https://link.aps.org/doi/10.1103/PhysRevLett.117.010503 (2016).

20. Granade, C. E., Ferrie, C., Wiebe, N. & Cory, D. G. Robust online Hamiltonian learning. *New Journal of Physics* **14,** 103013 (2012).

21. Knill, E., Laflamme, R. & Milburn, G. J. A scheme for efficient quantum computation with linear optics. *Nature* **409,** 46 (2001).

22. Ralph, T. C., Langford, N. K., Bell, T. B. & White, A. G. Linear optical controlled-NOT gate in the coincidence basis. *Phys. Rev. A* **65,** 062324. https://link.aps.org/doi/10.1103/PhysRevA.65.062324 (2002).

23. O'Brien, J. L., Pryde, G. J., White, A. G., Ralph, T. C. & Branning, D. Demonstration of an all-optical quantum controlled-NOT gate. *Nature* **426,** 264 (2003).

24. Zhou, X.-Q., Ralph, T. C., Kalasuwan, P., Zhang, M., Peruzzo, A., Lanyon, B. P. & O'Brien, J. L. Adding control to arbitrary unknown quantum operations. *Nature communications* **2,** 413 (2011).

25. Silverstone, J. W., Bonneau, D., Ohira, K., Suzuki, N., Yoshida, H., Iizuka, N., Ezaki, M., Natarajan, C. M., Tanner, M. G., Hadfield, R. H., *et al.* On-chip quantum interference between silicon photon-pair sources. *Nature Photonics* **8,** 104 (2014).

26. Zhou, X.-Q., Kalasuwan, P., Ralph, T. C. & O'Brien, J. L. Calculating unknown eigenvalues with a quantum algorithm. *Nature photonics* **7,** 223 (2013).

27. O'Malley, P. J. J. *et al.* Scalable Quantum Simulation of Molecular Energies. *Phys. Rev. X* **6,** 031007. https://link.aps.org/doi/10.1103/PhysRevX.6.031007 (2016).

28. Santagati, R., Gentile, A. A., Knauer, S., Schmitt, S., Paesani, S., Granade, C., Wiebe, N., Osterkamp, C., McGuinness, L. P., Wang, J., Thompson, M. G., Rarity, J. G., Jelezko, F. & Laing, A. Magnetic-Field Learning Using a Single Electronic Spin in Diamond with One-Photon Readout at Room Temperature. *Phys. Rev. X* **9,** 021019. https://link.aps.org/doi/10.1103/PhysRevX.9.021019 (2 2019).

29. McClean, J. R., Babbush, R., Love, P. J. & Aspuru-Guzik, A. Exploiting locality in quantum computation for quantum chemistry. *The journal of physical chemistry letters* **5,** 4368–4380 (2014).

30. McClean, J. R., Romero, J., Babbush, R. & Aspuru-Guzik, A. The theory of variational hybrid quantum-classical algorithms. *New Journal of Physics* **18,** 023023 (2016).

31. Wecker, D., Hastings, M. B. & Troyer, M. Progress towards practical quantum variational algorithms. *Phys. Rev. A* **92,** 042303. https://link.aps.org/doi/10.1103/PhysRevA.92.042303 (2015).

32. Cao, Y., Romero, J., Olson, J. P., Degroote, M., Johnson, P. D., Kieferová, M., Kivlichan, I. D., Menke, T., Peropadre, B., Sawaya, N. P., *et al.* Quantum Chemistry in the Age of Quantum Computing. *arXiv:1812.09976* (2018).

33. Cipra, B. A. The Ising model is NP-complete. *SIAM News* **33,** 1–3.

34. Lidar, D. A. & Wang, H. Calculating the thermal rate constant with exponential speedup on a quantum computer. *Phys. Rev. E* **59,** 2429–2438. https://link.aps.org/doi/10.1103/PhysRevE.59.2429 (1999).

35. Serrano-Andrés, L. & Merchán, M. Quantum chemistry of the excited state: 2005 overview. *Journal of Molecular Structure* **729,** 99–108 (2005).

36. Olsson, M. H., Mavri, J. & Warshel, A. Transition state theory can be used in studies of enzyme catalysis: lessons from simulations of tunnelling and dynamical effects in lipoxygenase and other systems. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* **361,** 1417–1432 (2006).

37. Crawford, T. D. & Stanton, J. F. Some surprising failures of Brueckner coupled cluster theory. *The Journal of Chemical Physics* **112,** 7873–7879 (2000).

38. Wang, L.-W. & Zunger, A. Solving Schrödinger's equation around a desired energy: application to silicon quantum dots. *The Journal of Chemical Physics* **100,** 2394–2397 (1994).

39. Colless, J. I., Ramasesh, V. V., Dahlen, D., Blok, M. S., Kimchi-Schwartz, M. E., McClean, J. R., Carter, J., de Jong, W. A. & Siddiqi, I. Computation of Molecular Spectra on a Quantum Processor with an Error-Resilient Algorithm. *Phys. Rev. X* **8,** 011021. https://link.aps.org/doi/10.1103/PhysRevX.8.011021 (2018).

40. Chandler, D. *Introduction to modern statistical mechanics* (Oxford University Press, 1987).

41. McClean, J. R., Kimchi-Schwartz, M. E., Carter, J. & de Jong, W. A. Hybrid quantum-classical hierarchy for mitigation of decoherence and determination of excited states. *Phys. Rev. A* **95,** 042308. https://link.aps.org/doi/10.1103/PhysRevA.95.042308 (2017).

42. Poulin, D., Hastings, M. B., Wecker, D., Wiebe, N., Doberty, A. C. & Troyer, M. The trotter step size required for accurate quantum simulation of quantum Chemistry. *Quantum Information & Computation* **15,** 361–384 (2015).

43. Fassioli, F., Dinshaw, R., Arpin, P. C. & Scholes, G. D. Photosynthetic light harvesting: excitons and coherence. *Journal of The Royal Society Interface* **11,** 20130901 (2014).

7 QUANTUM HAMILTONIAN LEARNING

In this chapter we will discuss a quantum application where instead of calculating properties of a well specified physical system, as in quantum simulation, we are interested in learning the physics underlying it. In particular, in all quantum simulation protocols discussed in previous chapters the problem we had to solve was quite specific. Namely, given the Hamiltonian of a system, infer some properties of it, e.g. its spectrum, its induced dynamics, or the transition probabilities when the Hamiltonian is perturbed. Whether or not the Hamiltonian used actually describes the real system is a different, maybe more fundamental, question. In fact, when we talk about simulations, we don't actually mean simulating the actual physical system, but rather our description of it, e.g. through its Hamiltonian model[1]. However, if we don't have an exact knowledge of the underlying physics of the system, or our approximations are too rough, our model may not be reliable. So the question is how to determine if our model of a quantum system is good or not, and, if it is not, how to efficiently learn a better description of it. In this chapter we will describe algorithms to address these issues.

A first interesting point to make is that, if we need to assess the fitness of a model, we will necessarily need to make predictions for it and compare them with the actual data obtained from the real system. In some sense, this is precisely what science is about: propose a theory (or model) and test if its predictions are consistent with the actual observation in Nature [3]. However, an important assumption here is that we are actually able to make predictions on the model. In other words, we have to assume that a model is simulatable if we want to test it. On the other hand, if we are not able to make any relevant predictions about a model, perhaps because it is too computationally expensive, there is no way to understand if the model itself is correct or not. Therefore, if a model

1 In fact, the simulated model can also be not physical at all! For example, the Dirac equation does allow unphysical solutions which have been simulated on a quantum processor [2]. Our physical simulator doesn't care if what it is simulating actually respects the laws of Nature or not.

cannot be efficiently simulated, we will no longer be able to assess whether it can reliably describe systems with sufficient complexity. The efficient quantum simulation of quantum models, generally intractable for classical computers, thus also represents an important tool for our understanding of fundamental physics.

In the context of quantum technologies, assessing the fitness of a Hamiltonian model and finding the Hamiltonian parameters which best describe a physical quantum system can be regarded as a verification and characterisation task for quantum hardware. The problem of finding classical methods to efficiently validate and characterise unknown quantum processes has been extensively studied, and is known to not be possible in general (see e.g. [4, 5]). In this chapter we will instead focus on using quantum simulators to validate and characterise unknown quantum devices.

To do it, we will describe an approach called *quantum Hamiltonian learning* (QHL), originally proposed by N. Wiebe, C. Granade, and others [6, 7]. The general idea of the algorithms is to combine quantum simulation with Bayesian inference methods to learn the Hamiltonian parameters that provide the best description of the dynamics of a unknown quantum system. After describing the algorithms, we will report an experimental implementation of QHL, where an integrated photonics quantum simulator is used to learn the Hamiltonian of an electronic spin system in a diamond NV centre. We will also present ideas to generalise QHL algorithms so that, instead of just learning the Hamiltonian parameters for a fixed Hamiltonian model, we can think of learning the underlying physical model itself.

7.1 Learning quantum Hamiltonians via quantum simulators

In this section we will describe the theory for the learning algorithms we will be using in the rest of the chapter. In general, they borrow ideas from machine-learning approaches, in particular Bayesian methods, and combine them with the capabilities of quantum simulators. The underlying learning algorithm drives experiments performed on both the unknown quantum system under investigation and the controllable quantum processor simulating the tested models. Quantum simulation will thus be used as a subroutine of the whole algorithm. We will touch on three types of cases, with different computational efficiencies of the learning algorithm: i) using classical machines for the simulation, ii) using quantum simulators, but without access to trusted quantum channels, iii) using both quantum simulators and trusted quantum channels.

7.1.1 *Bayesian approach to Hamiltonian learning*

Before proceeding, it is useful to specify the context and the problem in more detail. Let us consider a class of systems whose dynamics are generated by a Hamiltonian described by a model $H(\boldsymbol{x})$, where $\boldsymbol{x}$ is a real vector of parameters that determine each system. For example, the parameters can represent the interaction strengths in an Ising Hamiltonian $H(\boldsymbol{x}) = \sum_{(i,j)} x_{i,j} \hat{\sigma}_z^i \hat{\sigma}_z^j$. More generally, one can simply think of them as knobs that we can turn in our quantum simulator to control it, for which we want to find the optimal configuration to simulate a target unknown system. Hamiltonian learning can then be stated in the following way:

Hamiltonian learning problem. *Given an unknown or untrusted quantum system, and a Hamiltonian model $H(\boldsymbol{x})$ described by a set of parameters $\boldsymbol{x}$, find the parameter values $\boldsymbol{x}_0$ such that $H(\boldsymbol{x}_0)$ provides the best description of the quantum dynamics of the system within the model $H(\cdot)$.*

A point that remains to be clarified is what it means for a set of parameters to provide a "better description" of the system's dynamics. Intuitively, it would be the set which provides the predictions that best fit the data observed from the system. This concept can be formalised in a Bayesian scenario as follows. Suppose that we have some control parameters, denoted with c, to decide the different experiments that can be performed on the quantum system. These can involve, for example, the state $|\psi\rangle$ on which we initialise the system, the time t we let the system evolve, or the measurement we perform on the system. At the end of each experiment the measurement provides us a datum d. Suppose we perform N experiments where at the i-th experiment ($i \in \{1, \dots, N\}$) we set the control parameters c_i and obtain the datum d_i. After all experiments are performed we then have our set of data $\mathcal{D} = \{d_i\}$ (i.e. our experience) associated to the set of experiments performed $\mathcal{C} = \{c_i\}$. Based on the experiments performed and on the experience gathered, we build up information about the parameters, which, in a Bayesian approach, is encoded in the posterior distribution $p(\boldsymbol{x}|\mathcal{D};\mathcal{C})$. As already discussed in section 6.3, the posterior $p(\boldsymbol{x}|\mathcal{D};\mathcal{C})$ represents our confidence that the set of parameters $\boldsymbol{x}$ provides the correct description of the system, based on our experience $\mathcal{D}$ and the experiments performed $\mathcal{C}$. Our best guess is then provided by the parameters

$$\boldsymbol{x}_0 = \operatorname*{argmax}_{\boldsymbol{x}} p(\boldsymbol{x}|\mathcal{D};\mathcal{C}). \tag{7.1.1}$$

To estimate the distribution $p(\boldsymbol{x}|\mathcal{D};\mathcal{C})$ we can, once again, make use of Bayes' theorem:

$$p(\boldsymbol{x}|\mathcal{D};\mathcal{C}) = \frac{p(\mathcal{D}|\boldsymbol{x};\mathcal{C})p_0(\boldsymbol{x})}{\int p(\mathcal{D}|\boldsymbol{x'};\mathcal{C})p_0(\boldsymbol{x'})d\boldsymbol{x'}} \tag{7.1.2}$$

$$= \frac{\left[\prod_{i=1}^{M} p(d_i|\boldsymbol{x};c_i)\right] p_0(\boldsymbol{x})}{\int \left[\prod_{i=1}^{M} p(d_i|\boldsymbol{x'};c_i)\right] p_0(\boldsymbol{x'})d\boldsymbol{x'}}, \tag{7.1.3}$$

where we have assumed the independence of different experiments, and the initial prior distribution to be causally independent from $\mathcal{C}$ (i.e. the initial knowledge does not depend on the type of experiments one is going to perform). The quantity $p(d|\boldsymbol{x};c)$ is the likelihood function (which we already discussed in section 6.3), that represents the probability of obtaining the datum d from the experiment c according to the parameters $\boldsymbol{x}$. In other words, it quantifies how much the predictions according to $\boldsymbol{x}$ are consistent with actual observations. Crucially, in order to estimate it, the experimenter needs to be able to make such predictions. In particular, the estimation requires simulating the experiment c a sufficiently large number of times according to parameters $\boldsymbol{x}$ on the simulator, with $p(d|\boldsymbol{x};c)$ given by the fraction of times the outcome d occurs. To obtain the likelihood with a precision ϵ, $\mathcal{O}(1/\epsilon^2)$ repetitions of the experiment have to be performed, which is efficient if $p(d|\boldsymbol{x};c)$ is only polynomially small with the system size.

7.1.2 *Quantum Likelihood Estimation*

Having set a general context for Hamiltonian learning in a Bayesian framework, we proceed describing the different protocols that have been developed to perform it. In general, the approaches consist in hybrid classical-quantum algorithms where learning routines are run on a classical processor, which, in turn, drives subroutines performed on quantum hardware. What distinguishes the various protocols is the type of experiments performed, the design of the control parameters, and the quantum resources used.

First we can consider the simplest situation, shown in Fig. 7.1, where the only quantum resource we require, other than the unknown system, is a quantum simulator able to implement the model $H(\boldsymbol{x})$. The approach is called *quantum likelihood estimation* (QLE). In this case, the experiment design is as follows: we choose an initial state $|\psi\rangle$, evolve it for a time t, and measure the output in a basis that includes the state $|\psi\rangle$. The control parameters for the experiments are thus $c = \{|\phi\rangle, t\}$, and we can define a binary outcome of the experiment, given by the projective operators $\{|\psi\rangle\langle\psi|, \mathbb{1} - |\psi\rangle\langle\psi|\}$, where the output is $d = 0$ if the states collapses into $|\psi\rangle$, and $d = 1$ otherwise. The experiments are then simulated efficiently on the quantum simulator and repeated multiples times, and with different Hamiltonians $H(\boldsymbol{x})$, to estimate the likelihood function $p(d|\boldsymbol{x};c)$, which is

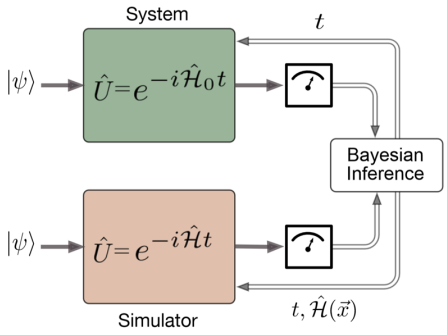


Figure 7.1: Schematic of QLE experiment designs. The system Hamiltonian H_0 (shaded green) is to be learned by a quantum simulator (shaded red) that embeds an abstract model $H(\boldsymbol{x})$ of the target system. The control parameters used in the experiment design are the initial state $|\psi\rangle$ and the evolution time t. The measurement projects the output state in the initial state ($|\psi\rangle\langle\psi|$) or in its orthogonal subspace ($1-|\psi\rangle\langle\psi|$). In the system, the initial state $|\psi\rangle$ is evolved for a time t and measured. The simulator mimics the system dynamics according to the model and obtains an estimate of the likelihood function from the statistics of the outcomes. The likelihoods are then used to infer the posterior distribution of the parameters $\boldsymbol{x}$ via Bayes' rule and to calculate the next step time t.

now of the form

$$p(0|\boldsymbol{x};c) = \left|\langle\psi|\, e^{-iH(\boldsymbol{x})t}\, |\psi\rangle\right|^2, \tag{7.1.4}$$

$$p(1|\boldsymbol{x};c) = 1 - p(0|\boldsymbol{x};c). \tag{7.1.5}$$

The measured likelihoods are then fed to an external classical processor which drives the Bayesian inference of the Hamiltonian parameters and proposes new experiments, as will be described in section 7.1.4.

Note that, instead of using a quantum simulator, we could think of using a classical computer to perform the subroutine involving the simulation of the evolution $e^{-iH(\boldsymbol{x})t}\,|\psi\rangle$ and the calculation of the likelihoods; an approach called *Classical Likelihood Estimation* (CLE) [6]. However, since classical simulation of quantum systems is strongly believed to be difficult in general, in this approach the algorithm cost is dominated by the simulation part, which is avoided in QLE by using a quantum simulator.

On the other hand, while QLE is in principle scalable, a challenge to the efficiency of the inference process is encountered for long evolution times, which is due to chaos emerging in the Hamiltonian evolutions [7]. It is well known that two nearly identical Hamiltonians will typically generate evolutions that diverge exponentially after a short time, before saturating at an exponentially small overlap; a famous effect in chaos theory knows as *Loschmidt echo* [8]. In particular, if two N-dimensional Hamiltonians H_a and H_b differ by an amount which is large compared to their characteristic spectral gaps, then, under realistic assumptions[2] there exist times $\tilde{t}$ and t_{eq} such that

2 An assumption for quantum chaos to arise is that the Hamiltonians are non-trivial (e.g. they are not

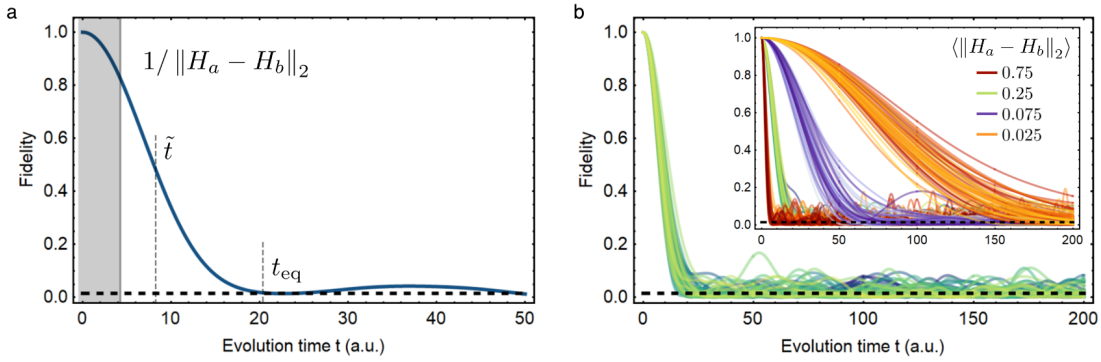


Figure 7.2: Chaotic behaviour for random Hamiltonians due to the Loschmidt echo. **a,** Exemplary evolution of the fidelity $\left|\langle\psi| e^{iH_b t} e^{-iH_a t} |\psi\rangle\right|^2$ of two systems with the same initial state but evolving with two slightly different Hamiltonians. The test is performed on a 6-qubit system, such that the dimension of the Hilbert space is $N = 64$. The distance between the two Hamiltonians was $\|H_a - H_b\|_2 \simeq 0.25$. The fidelity remains $\mathcal{O}(1)$ for small times $t \lesssim 1/\|H_a - H_b\|_2$, but a fast exponential decrease starts for evolution times $t \gtrsim \tilde{t}$, until a saturation at an exponentially small value $\simeq 1/N$ for times $t \gtrsim t_{\text{eq}}$. **b,** Same simulation as in a), but repeated with 50 random Hamiltonians. Approximately the same behaviour is observed in all cases. Inset: simulations with 50 repetitions for random Hamiltonians with different distances $\|H_a - H_b\|_2$.

$$\left|\langle\psi| e^{iH_b t} e^{-iH_a t} |\psi\rangle\right|^2 = \begin{cases} 1 - \mathcal{O}(t^2) & \text{if} \quad t \leq \tilde{t}, \\ e^{-\mathcal{O}(t)} & \text{if} \quad \tilde{t} \leq t \leq t_{\text{eq}}, \\ \frac{1}{N} & \text{for most} \quad t \geq t_{\text{eq}}, \end{cases} \tag{7.1.6}$$

where, roughly speaking, $\tilde{t}$ and t_{eq} depend on how much the Hamiltonians differ. This behaviour can be observed from the plots in Fig. 7.2a. Now, what the Loschmidt echo implies is that, if QLE experiments (where the likelihood is obtained setting $H_b = \mathbb{1}$ in eq. (7.1.6)) are performed with evolution times $t \geq t_{\text{eq}}$ then all the likelihoods will, with high probability, be $\mathcal{O}(2^{-n})$. Here n is the number of qubits in the system (i.e. $N = 2^n$). However, as we saw before, the estimation of the likelihood requires to repeat the experiment $\mathcal{O}(1/\epsilon^2)$ on the quantum simulator to achieve a precision ϵ. Therefore, if we want to resolve likelihoods with a magnitude of $\mathcal{O}(2^{-n})$, the number of repetitions required explodes exponentially with the size of the system n, making the protocol not

both multiples of the identity, the input state $|\psi\rangle$ is not a common eigenstate, etc.). These conditions are typically true for most systems one wants to investigate with Hamiltonian learning, although relevant cases where they are not respected exist (see for example Ref. [9], where periodic, not chaotic, dynamics arise for an NV centre in diamond under the action of an external magnetic field). This assumption can be mathematically formalised by requiring that at least one of the two Hamiltonians has complexity that is characteristic of canonical random matrix ensembles, see Ref. [8] for more details.

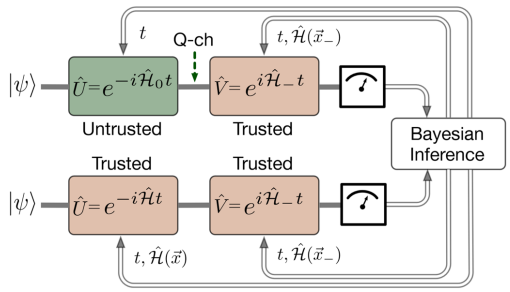


Figure 7.3: Schematic of IQLE experiment designs. The untrusted quantum system is shaded green and the trusted quantum devices are shaded red. IQLE lies in the inversion of the evolution via a Hamiltonian H_- implemented with a trusted device (top one). The trusted and the untrusted parts are linked by a coherent quantum channel. The inversion is performed also in the likelihood estimation on the trusted quantum devices (bottom). Results are classically processed for Bayesian inference.

scalable. To avoid this issue, we are forced to use small evolution times $t < t_{\text{eq}}$. On the other hand, experiments with small evolution times are less informative[3], and this limitation can thus significantly slow down the inference process.

Another type of experiment design is thus needed if we want a protocol which is both scalable and highly informative.

7.1.3 *Interactive Quantum Likelihood Estimation*

Before describing how better experimental designs can be engineered, it is useful to build a better intuition of where QLE fails. From the discussion above we can say that the issue arises due to a fast decay in the likelihood functions for long evolution times. In fact from 7.2b we can heuristically observe that the steep decay of the likelihood function, due to Loschmidt-echo-type effects, starts at evolution times $t \approx 1/\left\|H_a - H_b\right\|_2$ (an analytical argument for this behaviour can be found in Ref. [7, 8]). Now, to estimate the likelihood in a QLE experiment (eq. 7.1.4) we need to fix H_b to the identity in eq. (7.1.6), which then poses serious limits to the evolution time if we want to avoid exponentially small likelihoods, unless also H is trivially close to the identity.

To resolve this issue, an interactive approach can be used, called *interactive quantum likelihood estimation* (IQLE). The protocol is schematised in Fig. 7.3. After the evolution under H_0 is performed on the system (or under H on a simulator), the output state is now fed, via a coherent quantum channel, into a quantum simulator which reverses it according to an Hamiltonian H_-. In this case we thus have $H_b = H_-$, which is no longer fixed as in QLE. Now, the idea is that, while we are learning more information about the true Hamiltonian of the system H_0, we can adapt H_- iteratively from our knowledge of H_0. In this way, the difference between the two Hamiltonians $\left\|H_0 - H_-\right\|_2$ decreases as we proceed during the protocol, allowing to increase the evolution time

3 The intuition of experiments with small evolution times being uninformative can be formalised in terms of the Fisher information, which for QLE experiments scales as $\mathcal{O}(\|H\|^2 t^2)$ [10, 11].

$t \approx 1/\|H_0 - H_-\|_2$ to get more informative experiments while maintaining an efficient estimation of the likelihoods. In fact, with high probability, test Hamiltonians with $H \approx H_-$ will have efficiently calculable likelihoods of magnitude $\mathcal{O}(1)$. In other words, for these cases, even if the state is scrambled by the forward evolution, inverting under a well chosen Hamiltonian H_- will unscramble it, making the associated likelihood easy to estimate.

The IQLE protocol is thus both scalable, i.e. provides an efficient likelihood estimation, and highly informative. Additionally, it has been numerically shown to be more robust to noise, with respect to QLE [10]. This comes at the cost of requiring a quantum communication between the untrusted system and a trusted simulator that inverts the evolution on the output state received from the system (see Fig. 7.3). This can be performed via a `swap` gate or a direct coherent quantum channel, both of which are quite challenging to obtain in practice with current quantum hardware capabilities [7]. Nevertheless, an experimental photonic implementation of IQLE will be reported in section 7.2.5.

7.1.4 *Sequential Monte-Carlo algorithms for the Bayesian inference of Hamiltonian parameters*

Now that the QLE and IQLE protocols for the quantum Hamiltonian learning problem have been described, we can detail the methods, adapted from machine-learning applications, to perform the Bayesian inference process in a practical and efficient manner. As for the Bayesian phase estimation protocol discussed in section 6.3, the main practical problem we encounter when applying Bayes' rule to infer Bayesian parameters (as in eq. 7.1.3) is the calculation of the normalisation integral at the denominator, which would require the calculation of a continuous amount of likelihood functions. Similarly to section 6.3, we can avoid this problem by discretising the parameter space using a finite set of parameter values $\{\boldsymbol{x_k}\}$, called *particles*. However, we now need to generalise the methods used for the RFPE algorithm of section 6.3 in two ways: i) we will need to deal with more than a single parameter, and ii) the likelihood functions here are no longer periodic as in phase estimation, meaning that the choice of optimal experiments for the inference process will be different.

To make such a generalisation, we can adapt *sequential Monte-Carlo* (SMC) algorithms to quantum Hamiltonian learning [6, 7, 12]. These types of algorithms, also known as *particle filters*, are a class of approximate inference algorithms that are increasingly popular for approximated Bayesian inference. The central idea is to represent distributions in the

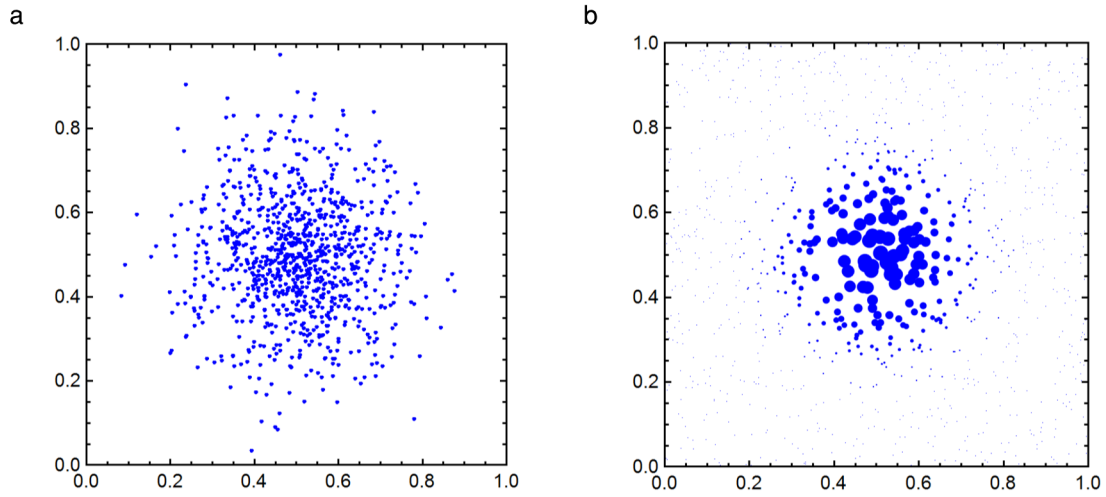


Figure 7.4: Two different ways of representing a multinormal probability distribution in the SMC approach. **a,** Using particle density, and setting the weights to be uniform. **b,** Using particle weights. Each point indicates a particle position, with the radius proportional to the associated particle weight.

parameter space as weighted sums of delta functions

$$p(\boldsymbol{x}) \approx \sum_{k=1}^{M} w_k \delta(\boldsymbol{x} - \boldsymbol{x}_k), \tag{7.1.7}$$

where $\{\boldsymbol{x}_k\}_{k=1}^M$ is a set of M particles, and $\{w_k\}_{k=1}^M$, with $\sum_k w_k = 1$, are the *weights* associated to the particles. With this representation, the parameter space is discretised, with the mean and covariance of the distribution readily obtainable as $\boldsymbol{\mu} = \sum_{k=1} w_k \boldsymbol{x}_k$ and $\Sigma = \sum_k w_k \boldsymbol{x}_k \boldsymbol{x}_k^\top - \boldsymbol{\mu}\boldsymbol{\mu}^\top$. Moreover, we have two ways to describe a particular distribution: keeping the particle weights constant and changing the position of the particles, or keeping the particles position fixed and adjusting the weights. These two situations are shown in Fig. 7.4. The Bayesian update of the distribution is performed in the second type of description. That is, upon the measurement of an outcome d from an experiment c, the particle positions are left unchanged, and the weights updated via the discretised version of eq. (7.1.3):

$$w_k \mapsto \frac{w_k \; p(d|\boldsymbol{x}_k; c)}{\sum_k w_k \; p(d|\boldsymbol{x}_k; c)}, \tag{7.1.8}$$

with $p(d|\boldsymbol{x}_k; c)$ the likelihoods which can be measured via calling M estimations on the quantum simulator. However, similarly to section 6.3, if we never move the particles

as the inference proceeds, we will encounter a problem when the distribution will shrink to a point where only few particles will have non-negligible weights. In such situations the description of the distribution is no longer adequate. This issue can be formalised by defining the *effective particle number* as the inverse participation ratio $M_{\text{eff}} = 1/\sum_k w_k^2$, which quantifies how many particles are effectively contributing in describing the distribution [6]. Since at each update the particle weights distribution typically gets more peaked, M_{eff} drops as more experiments are performed, indicating a worse description of the distribution. To avoid it, we can define a threshold for $\tilde{M}_{\text{eff}}$, usually $\tilde{M}_{\text{eff}}/M = 0.5$, below which the particles are resampled according to the current description of the prior distribution, and their weights reinitialised to uniform ($w_k = 1/M$). There are may different approaches that can be used to choose the new particle positions when resampling in SMC algorithms. In the context of Hamiltonian learning, the typical choice is the Liu-West resampler, which has shown particularly good numerical stability in absence of degeneracies in the model[4] [6].

Although SMC can be used in an *offline* manner, where the inference is performed after all measurements have been performed, a more efficient way to do it is performing the learning *online*. In the latter case, the experiments are performed iteratively, where at each step the control parameters c for the next experiment are chosen, based on the prior distribution, to maximise its informativeness.

Here we will focus on online Hamiltonian learning, which, for both QLE and IQLE, proceeds as follows. We start with an initial prior distribution $p_0(\boldsymbol{x})$ in the parameter space, which represents our a-priori knowledge on the system. If no a-priori information is available, a conservative choice is a uniform distribution over an arbitrarily large section of the parameter space. We initialise the SMC algorithm by sampling M particles from the prior, and assign them uniform weights $w_k = 1/M$. We then start performing the experiments iteratively. At the i-th step of the algorithm, the control parameters c_i for the experiment are chosen based on the prior, and outcome d_i is obtained performing the experiment on the system. The experiment is then simulated on the quantum simulator to estimate the likelihoods $p(d_i|\boldsymbol{x}_k; c_i)$ for all M particles $\{\boldsymbol{x}_k\}$, and the posterior obtained updating the particle weights as in eq. (7.1.8). The effective particle number M_{eff} is then calculated, and if it is below the threshold $\tilde{M}_{\text{eff}}$, the resampling subroutine is performed, obtaining a new set of particles and reinitialising their weights to uniform. This algorithm then proceeds iteratively using the posterior as the new prior for the next iteration.

An important question is now how to choose the best control parameters c to optimise the learning rate of the Hamiltonian parameters in online experiments. In particular,

4 Very recently, new techniques combining SMC methods with k-clustering algorithms (a standard in machine learning) have been developed to tackle also the cases where degeneracies arise in inferring the model parameters in quantum Hamiltonian learning [13].

we are interested in the choice of the evolution time t for both QLE and IQLE, and the inversion Hamiltonian H_- for IQLE. The input state $|\psi\rangle$ will instead be considered fixed at the start for simplicity, and not regarded as a control parameter. Following the same spirit as in section 6.3, an adaptive heuristic strategy, known as the *particle guess heuristic* (PGH) can be employed to provide near-optimal experiments. Roughly speaking, this strategy, developed in Ref. [6, 7], takes the evolution time at each iteration step to scale as the reciprocal of the uncertainty in the current prior. In particular, for QLE experiments we will use $t = 1.26/\left\|\Sigma\right\|_2^{1/2}$, which for a single parameter Hamiltonian reduces to the same PGH $t = 1.26/\sigma$ used in the RFPE algorithm (see section 6.3) [6, 7, 14]. For IQLE experiments, we adopt instead the PGH $t = 1/\left\|H(\boldsymbol{x}_\times) - H(\boldsymbol{x}_-)\right\|_2$, where $\boldsymbol{x}_\times$ and $\boldsymbol{x}_-$ are sampled from the current prior, and the inversion Hamiltonian given by $H_- = H(\boldsymbol{x}_-)$. In both cases, these heuristics are known to provide near-optimal experiments for Hamiltonian learning [6, 7, 14].

7.2 Experimental implementation of quantum Hamiltonian learning

We will now describe a demonstration of the quantum Hamiltonian learning algorithms described above. To implement the protocols, we needed to have access to both a system to be tested and a quantum simulator. In particular, to better manifest the universality of the algorithms, it was interesting to use a quantum system and a simulator embedded in different platforms and described by different physics. In this case, when performing the inference, we are really inducing a quantum system to mimic the behaviour of another one even if the physics underlying them are completely different, by exploiting the ability of one to simulate the other.

To implement QLE, we used an electron spin in a negatively charged nitrogen-vacancy (NV^-) centre in diamond as an untrusted system [15, 16], and a silicon quantum photonic device as a quantum simulator. The two systems were interfaced together via a classical processor driving both setups and performing the Bayesian inference algorithm, as shown in Fig. 7.5. For IQLE, given the requirement of a quantum channel, currently impractical to implement between an NV centre which emits at visible wavelengths and a silicon chip operating at telecom wavelengths, we decided to have both the system and the simulators on the same chip. Doing so, we could straightforwardly implement a coherent quantum channel via integrated photonic waveguides, making IQLE feasible.

We will start with a description of the quantum systems used in the experiment, and proceed reporting the experimental results.

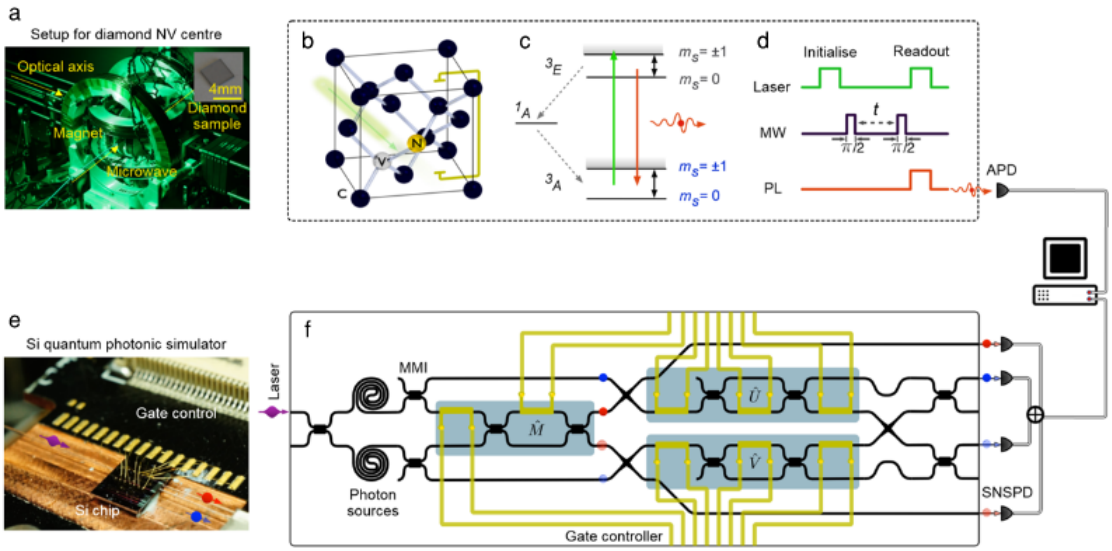


Figure 7.5: Setup for quantum Hamiltonian learning. **a,** Confocal setup with diamond (inset) containing NV^- centres. **b,** Structure and **c,** energy level diagram of an NV^- centre in diamond. The ground-state electron spin Hamiltonian, describing the coherent dynamics between $m_s = 0$ and -1, is to be characterised and learned using the quantum simulator. **d,** A single Rabi sequence for the initialisation, manipulation and read-out of the electron spin state. A laser pulse is used to initialise the spin into $m_s = 0$. Two microwave (MW) $\pi/2$-pulses with a time t delay are then used to coherently drive the spin. The spin state is measured by detecting photo-luminescence (PL) with an avalanche photodiode (APD). **e,** Silicon chip embedding the quantum photonic simulator. The device schematic is reported in **f**, where black lines are silicon nano-photonic waveguides, and gold wires represent thermo-optical phase-shifters and their transmission lines. A pair of photons with different wavelengths are generated via spontaneous four-wave mixing in the spiral waveguide sources. The operations $\hat{U}$ or $\hat{V}$ performed on the target photon are coherently controlled by the state of the other photon. Performing measurements $\hat{M}$ on the control photon allows to estimate the likelihood function for the chosen configuration of the device. The two different physical systems shown in **a)** and **e)** are interfaced through a classical computer.

7.2.1 *Quantum system: NV^- centre in diamond*

As untrusted quantum system for QLE, we employed a single solid-state spin-qubit encoded in a diamond NV^- centre. The NV colour centre in diamond consists of a nitrogen impurity adjacent to a vacancy in the diamond lattice (Fig. 7.5b). We will consider here the charged NV^- centre, which posses a triplet ground state (^3A level in Fig. 7.5c) with an electron angular momentum $S = 1$. Furthermore, in the ground state the sub-levels are split in energy by spin-spin interaction into a singlet state of spin projection $m_s = 0$ and a doublet $m_s = \pm 1$ [17]. The splitting of these ground state sub-levels is approximately 2.9 GHz, which means that the ground state transition is

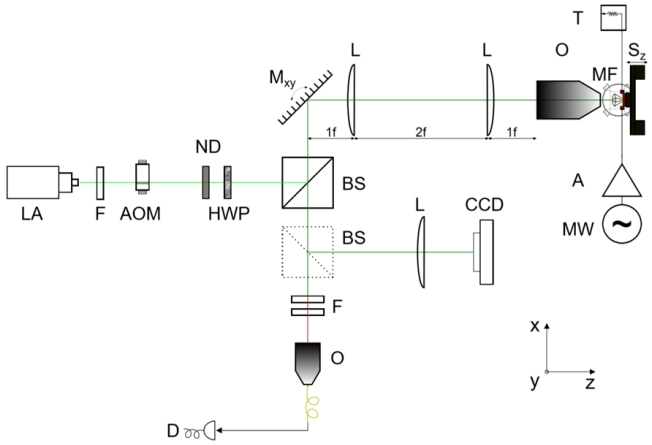


Figure 7.6: Schematic of the confocal setup used to control the the electron spin system in the diamond NV^- centre. For the off-resonant excitation of the NV^- centre a 532 nm CW laser (LA) is used. A filter (F) is placed after the laser to remove spurious side-bands. The excitation beam is then coupled in and out of an acousto-optic modulator (AOM) to to perform frequency scans. The laser power is controlled with neutral density (ND) filters and the laser polarisation with an additional half-wave plate (HWP). The laser is reflected on a beam splitter (BS) towards a fast steering mirror ($M_{x,y}$). The laser is then coupled through two lenses onto the backplane of an objective (O), and focused on the sample. The fluorescence and reflection from the sample passes back through the same optical path. At the BS, the emitted light passes through and enters the coupling objective. Filters (F) are used to eliminate spurious reflected laser light and Raman signals, allowing the detection of the emitted fluorescence via an avalanche photo diode detector (D). For the precise alignment of the magnetic fields to the axis of the NV^- centre, the sample is located at the centre of three current-controlled Helmholtz coil pairs (MF) (also shown in Fig. 7.5a). A microwave source (MW) signal generator is used for the programming of pulse sequences to manipulate the electron spin in the NV^- centre.

accessible with microwave excitation. The energy degeneracy in the $m_s = \pm 1$ levels can be lifted by applying an external magnetic field, which induces energy shifts of opposite sign to the $m_s = +1$ and $m_s = -1$ states via the Zeeman effect. By applying a magnetic field of sufficient magnitude (typically a few mT is enough) the two levels are separated enough so that the transition $m_s = 0$ to $m_s = -1$ can be individually addressed using the microwave signal, forming a two-level controllable system. This platform can be used to encode and control a qubit, where the computational basis is given by $|0\rangle = |m_s = 0\rangle$ and $|1\rangle = |m_s = -1\rangle$. Initialisation and readout of the qubit can be performed optically by exploiting the optical transition from the ground into the excited state (^3E in Fig. 7.5c), which also forms a triplet. In fact, this optical transition, with a resonance at 637 nm, is spin-preserving, meaning that the information encoded in the qubit is maintained. However, when decaying back from the excited state to the ground state different transition pathways are possible depending on the spin state.

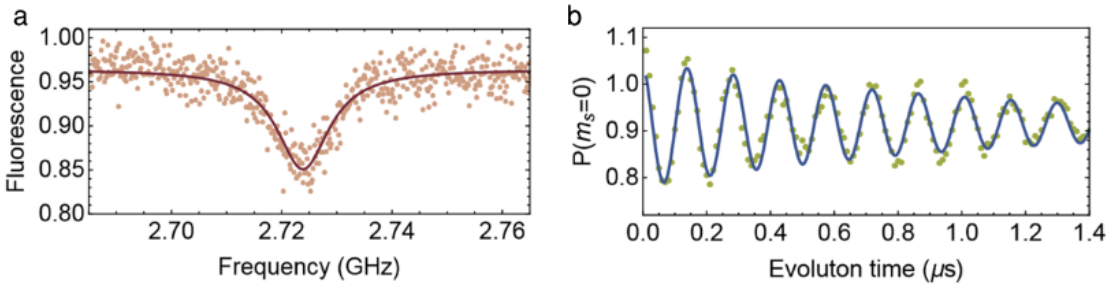


Figure 7.7: Preliminary characterisation of the NV^- centre. **a,** Optical detected magnetic resonance of the $m_s = 0$ to $m_s = -1$ transition. A central frequency of 2.742 GHz is measured. Orange points are data and red curve is a Lorentian fit of the resonance. **b,** Exemplary Rabi oscillation. Each data (green points) consists of a single Rabi sequence with three million iterations to gather sufficient statistics. The blue curve is a sinusoidal fit which includes decoherence.

In particular, while both $m_s = 0$ and $m_s = -1$ states can decay back to the ground state (preserving the spin) emitting a single-photon at 637 nm, the $m_s = -1$ state has $\approx 20\%$ higher chance to undergo a phonon-mediated non-radiative decay (i.e. no photons emitted) which brings it to $m_s = 0$ in the ground state. By repeating experiments many times and measuring the fluorescence emitted by the NV^- centre, i.e. the number of photons emitted in the decay, it is possible to infer the probabilities $p(m_s = 0)$ and $p(m_s = -1)$. This provides a measurement in the computational basis of the qubit. Moreover, as the presence of the non-radiative decays imply that the spin state has higher probability to be in the $m_s = 0$ every time the excitation and decay of the spin is performed, the qubit can be initialised in the $|0\rangle = |m_s = 0\rangle$ by repeating this procedure a sufficiently large number of times.

The setup used to control the NV^- centre in our experiment consisted of a home-built scanning confocal microscope, shown in Fig. 7.6. For initialisation and measurement of the electron spin, the NV^- centre is excited non-resonantly using a 532 nm CW laser, with the near-637 nm emitted fluorescence collected by an avalanche photo-diode (APD) single photon detector. To lift the degeneracy between the $m_s = \pm 1$ state, a 5 mT magnetic field was precisely aligned to the NV^- centre axis via three current-controlled Helmholtz coil pairs. In this way the resonance between the $m_s = 0$ and $m_s = -1$ states could be individually driven by 2.742 GHz microwave pulses, as shown in Fig. 7.7a, allowing the control and manipulation of the spin qubit.

The dynamics we are interested in studying to perform quantum Hamiltonian learning are the solid-state spin-qubit dynamics between the $m_s = 0$ and $m_s = -1$ states of the electron ground-state triplet in the NV^- centre, i.e. Rabi dynamics [15, 16]. In our system, these can be modelled by a single parameter Hamiltonian of the form $\hat{H}(f) = \hat{\sigma}_z f/2$

acting on the initial state, where $\hat{\sigma}_z$ is the Pauli z operator and the Hamiltonian parameter f is the *Rabi frequency*. The experiment we perform to implement quantum Hamiltonian learning can be described as:

1. The spin is optically initialised in the $|0\rangle = |m_s = 0\rangle$ state.
2. A $\pi/2$ microwave pulse is sent, preparing the system in the state $|+\rangle = (|m_s = 0\rangle + |m_s = -1\rangle)/2$.
3. The system performs a free evolution for a time t.
4. A second $\pi/2$ microwave pulse is sent.
5. The system is optically readout in the computational basis.

This sequence produces Rabi oscillations, which represent the type of experiments needed for QLE with $|\psi\rangle = |+\rangle$. To obtain a single-shot outcome d, required for QHL, the probability $p(d)$, where we define $d = 0$ ($d = 1$) if the state is projected in $|+\rangle$ ($|-\rangle$), is estimated from repeated experiments (three million iterations, leading to approximately 3500 photon counts), and the outcome d is determined using a majority voting scheme as in section 6.4. The oscillatory behaviour of the Rabi evolutions described above is shown in Fig. 7.7b, where the fluorescence signal, from which the probability $p(0)$ can be obtained, is reported for different evolution times.

7.2.2 *Quantum simulator: silicon photonic chip*

For the quantum simulator we used a silicon quantum photonics system, shown in 7.5e. The set-up and the silicon quantum photonic device used are the same ones employed for the experiments in section 6.4 and section 6.7, described in details therein. In those experiments the device was used to implement arbitrary control-unitary operations between two qubits. Here, we are going to use some slight modifications. In fact, as can be observed from the circuit in Fig. 7.5f, we use two more thermal phase-shifters on the target photon, which where set to identity in previous experiments, to implement an additional arbitrary unitary. In this way, proceeding in the same way as in section 6.4, the circuits implements a two-qubit gate where two unitaries $\hat{U}$ and $\hat{V}$ can be applied on the target qubit conditional on the state of the control one, producing the state

$$\frac{|0\rangle_1 \hat{U} |\psi\rangle_2 + |1\rangle_1 \hat{V} |\psi\rangle_2}{\sqrt{2}}, \tag{7.2.1}$$

where $\hat{U}$ and $\hat{V}$ and the input state $|\psi\rangle$ can be chosen arbitrarily by reconfiguring the integrated phase shifters.

To use this chip as a quantum simulator for QHL experiments we need to be able to calculate the likelihood functions for the experiment performed on the NV^- centre, i.e. simulating the evolution according to the Rabi Hamiltonian $\hat{H}(f)$ for arbitrary values of the parameter f. Note that the Rabi Hamiltonian is a single qubit Hamiltonian, while in the photonic system we have two qubits. This adds a bit of redundancy in the simulator, which we can here use to calculate the likelihoods using an entanglement-based scheme in the following manner. Starting from the output state in eq. (7.2.1), if we don't measure the second qubit we obtain for the first qubit the reduced density matrix

$$\begin{aligned}
\rho_1 &= \mathrm{Tr}_2(\rho_{1,2}) && (7.2.2)\\
&= \left[|0\rangle\langle 0| + |1\rangle\langle 1| + |0\rangle\langle 1|\,\mathrm{Tr}\left(\hat{U}|\psi\rangle\langle\psi|\hat{V}^\dagger\right) + |1\rangle\langle 0|\,\mathrm{Tr}\left(\hat{V}|\psi\rangle\langle\psi|\hat{U}^\dagger\right)\right]/2 && (7.2.3)\\
&= \Big[|+\rangle\langle +|\left(1+\mathrm{Re}\left(\langle\psi|\hat{U}^\dagger\hat{V}|\psi\rangle\right)\right) + |-\rangle\langle -|\left(1-\mathrm{Re}\left(\langle\psi|\hat{U}^\dagger\hat{V}|\psi\rangle\right)\right)\\
&\quad + i\left(|+\rangle\langle -| - |-\rangle\langle +|\right)\mathrm{Im}\left(\langle\psi|\hat{U}^\dagger\hat{V}|\psi\rangle\right)\Big]/2 && (7.2.4)\\
&= \Big[|+i\rangle\langle +i|\left(1+\mathrm{Im}\left(\langle\psi|\hat{U}^\dagger\hat{V}|\psi\rangle\right)\right) + |-i\rangle\langle -i|\left(1-\mathrm{Im}\left(\langle\psi|\hat{U}^\dagger\hat{V}|\psi\rangle\right)\right)\\
&\quad + \left(|-i\rangle\langle +i| - |+i\rangle\langle -i|\right)\mathrm{Re}\left(\langle\psi|\hat{U}^\dagger\hat{V}|\psi\rangle\right)\Big]/2, && (7.2.5)
\end{aligned}$$

where $\{|+\rangle, |-\rangle\}$ ($\{|+i\rangle, |-i\rangle\}$) is the eigenbasis of the Pauli X (Y) operator. Therefore, from eq. (7.2.4), we have that, when measuring in the Pauli X basis, the probability p_+ of measuring the first photon in the state $|+\rangle$ is related to the real part of the scalar product $\langle\psi|\hat{U}^\dagger\hat{V}|\psi\rangle$ via

$$\mathrm{Re}\left(\langle\psi|\hat{U}^\dagger\hat{V}|\psi\rangle\right) = 2p_+ - 1. \tag{7.2.6}$$

Similarly, from eq. (7.2.4) we have that information about the imaginary part can be obtained measuring the first qubit in the Pauli Y basis:

$$\mathrm{Im}\left(\langle\psi|\hat{U}^\dagger\hat{V}|\psi\rangle\right) = 2p_{+i} - 1, \tag{7.2.7}$$

where p_{+i} is now the probability of projecting in $|+i\rangle$. These quantities can be straightforwardly used to measure the likelihoods for QHL experiments. In fact, recalling that for QLE experiments the likelihoods are given by $\mathcal{L}_{\mathrm{QLE}} = |\langle\psi|e^{-i\hat{H}(\boldsymbol{x})t}|\psi\rangle|^2$, one can easily notice that it can be obtained in the photonic simulator by setting $\hat{U} = \hat{\mathbb{1}}$ and $\hat{V} = e^{-i\hat{H}(\boldsymbol{x})t}$ and using eq. (7.2.6) and eq. (7.2.7):

$$\mathcal{L}_{\mathrm{QLE}} = (2p_+ - 1)^2 + (2p_{+i} - 1)^2. \tag{7.2.8}$$

For IQLE experiments, the likelihood is given by $\mathcal{L}_{\text{IQLE}} = |\langle\psi|e^{i\hat{H}(\boldsymbol{x}_-)t}e^{-i\hat{H}(\boldsymbol{x})t}|\psi\rangle|^2$, and can obtained in the same way as for eq. (7.2.8), but now setting $\hat{U} = e^{-i\hat{H}_-t}$ and $\hat{V} = e^{-i\hat{H}(\boldsymbol{x})t}$.

Note also that, using this entanglement-based scheme to implement IQLE, all evolutions are forward in time, in contrast with the original approach where the time reversal $e^{i\hat{H}_-t}$ is performed by a backwards evolution. This property can make the entanglement-based approach amenable for analog quantum simulators. However, it comes at the cost of additional quantum resources, i.e. entanglement between the system and an ancillary qubit.

7.2.3 *Experimental results for QLE*

To implement QLE, both the silicon-photonics device and the NV^- centre were interfaced via a classical computer, as shown in Fig. 7.5, which was driving both setups and performed the Bayesian updates. The goal is to learn the Hamiltonian acting on the initial state for the Rabi oscillations of the NV^- centre's electron spin, using the Rabi model $\hat{H}(f) = \hat{\sigma}_x f/2$. The Hamiltonian parameter to be learnt was thus the the Rabi frequency f. We use the capability of the the silicon-photonics chip to simulate the model $\hat{H}(f)$ for arbitrary f to calculate the likelihood function for each particle. At each step of the QLE implementation, the evolution time t was chosen adaptively according to the PGH. Photo-luminescence results from the NV^- centre were calculated from 3 million iterations for each sequence, and fed to the classical processor to extract a single shot outcome d using majority voting. The photonic simulator then repeated the experiment with the same evolution time t to obtain the likelihoods for each particle, which were then used to update the prior distribution via the classical computer. The algorithm proceeded by iteratively repeating these steps. The prior distribution $p(f)$ of the particles was initialised to be uniform between 0 MHz and an arbitrary large value Δf, for which we chose $\Delta f = 100/2\pi$ MHz. A total of 20 particles were used in the SMC algorithm, which, from preliminary simulations, resulted to be a sufficient number.

Experimental results for QLE are shown in Fig. 7.8 for 50 steps of the algorithm. For more clarity, all plots report the rescaled dimensionless quantity $\omega = f/\Delta f$, for which the prior is initially uniformly distributed in the interval $\omega \in [0, 1]$. Fig. 7.8a shows the particle's distribution converging to the correct value of the parameter ω_0. This can be also observed in Fig. 7.8b where the mean and standard deviation of the posterior distribution are reported for different steps of the QLE algorithm. The final learned value corresponds to the Rabi frequency $f_{\text{QLE}} = (6.93 \pm 0.09)$ MHz, given by the mean and standard deviation of the final posterior distribution after 50 steps. This result is consistent with the reference value $f_0 = 6.90$ MHz obtained with a direct fit

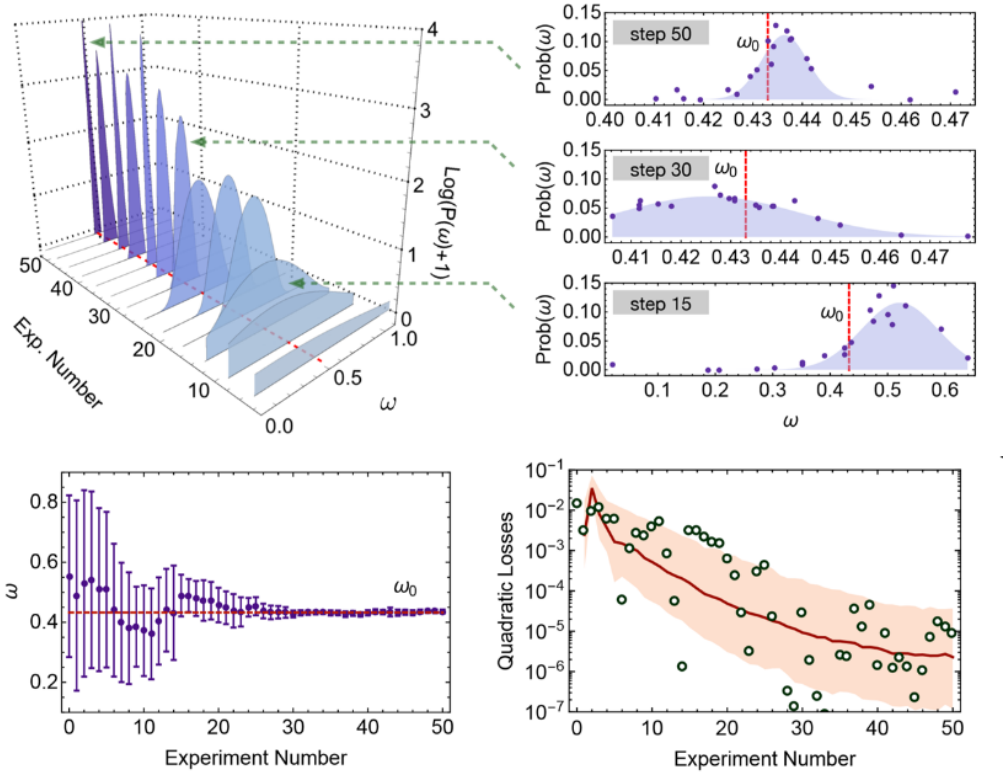


Figure 7.8: Experimental results for QLE. **a,** Progressive learning of the electron spin Hamiltonian, parametrised by a rescaled Rabi frequency $\omega = f/\Delta f$, via QLE. The distribution over the Hamiltonian parameter ω is described by a discrete approximation using 20 particles. Within 50 steps, the distribution converges to the correct value ω_0 (dashed red line). Insets: the distribution of particles after 15, 30 and 50 steps. Points represents particle positions and weights, and shaded areas are Gaussian fittings of the associated distribution. **b,** Evolution of the mean and standard deviation of the distribution. Error bars are ± 1 standard deviation. **c,** Evolution of the quadratic losses. Circles are experimental data and the line represents theoretical simulation results with a 67.5% confidence interval (shaded area). The theoretical simulation was averaged over 500 runs of QLE.

of the Rabi oscillations (see Fig. 7.7b). Therefore, the simulator successfully learns the parameter that best represents this Hamiltonian, without prior knowledge of the Rabi frequency. The fast experimental convergence of the algorithm to ω_0 can be also observed through the evolution of the quadratic losses of the particles distribution (here equal to the mean-squared error) achieving a final value of approximately 10^{-5} (see Fig. 7.8c).

7.2.4 *Beyond Hamiltonian learning: quantum model learning*

So far we have considered a learning task where the model of the Hamiltonian $H(\boldsymbol{x})$ was fixed and we looked for the optimal parameters $\boldsymbol{x}$ to best describe the system within the proposed model. However, as already discussed in the introduction to this chapter, a priori we don't know if the model used is appropriate or not. For example, this may happen if some of the approximation used in deriving it were not adequate, or because our understanding of the physics underlying the system was incorrect or incomplete. Such failures of the model can be identified by observing predictions that are inconsistent with the measured data. Interestingly, we were able to observe and quantify the inconsistencies of the Rabi model used to describe the spin dynamics while performing the QLE algorithm. These are manifested in the evolution of the posterior distribution variance for the parameter ω during the QLE algorithm, reported in Fig. 7.9. It can be observed that, while the variance decreases exponentially during, approximately, the first 35 steps, representing a rapid learning regime for the algorithm, the variance σ^2 saturates at $\bar{\sigma}^2_\omega \simeq 4.2 \times 10^{-5}$. Such saturation indicates a limit in the amount of information we can extract about the system, due to an incomplete description of the spin system via the model used.

Knowing when a model has failed affords us the opportunity to improve upon it. In fact, we can in principle use the information obtained from detecting and characterising such inconsistencies between our current model and the observed data to improve the quantum model itself. We named this problem *quantum model learning*, and tried to take first steps to address it in an algorithmic way. In particular, to improve our initial model using the Rabi Hamiltonian $\hat{H}(f) = \hat{\sigma}_x f/2$, which we will now call Model I, we can propose a new model (Model II) inserting a time dependent term in the Hamiltonian: $\hat{H}(f, \alpha; t) = \hat{\sigma}_x(f + \alpha t)/2$. This is now a two-parameter model, where we added a chirping

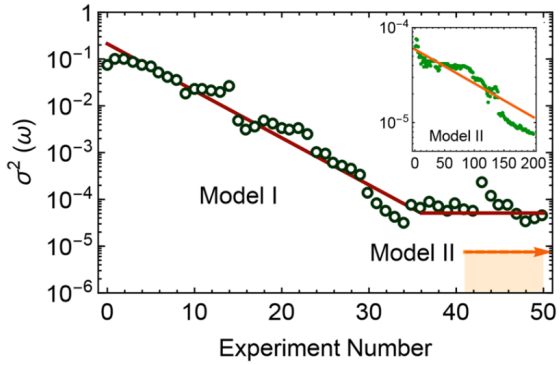


Figure 7.9: Model validation and improvement. The presence of other physical effects in the system that are not describable by the model $\hat{H}(\omega)$ (Model I) limits the amount of extractable information, as manifested by a saturation of the distribution variance at $\sigma^2_\omega \simeq 4.2 \times 10^{-5}$ after approximately 35 steps. The adoption of a new two-parameters model $\hat{H}'(\omega, \xi)$ (Model II), which includes the presence of chirping, allows to achieve a covariance below $\|\Sigma\|_2 = 7.5 \times 10^{-6}$ (the shaded area). Inset: covariance norm evolution of Model II.

factor α. We now need a procedure to test, given the final distribution of the parameter for Model I, whether Model II is a more accurate choice or not. The principle we adopt to develop such procedure is as follows.

First, we perform standard QHL also for Model II, using the QLE algorithm. The results for QLE on Model II are reported in Fig. 7.10. Similarly to the implementation of QHL for Model I, for simplicity we assume the parameters values within the intervals $f \in [0, \Delta f]$ and $\alpha \in [-\Delta\alpha/2, \Delta\alpha/2]$, with arbitrary large values $\Delta f = 100/2\pi$ MHz and $\Delta\alpha = 10^4/2\pi$ MHz^2, and consider the two rescaled dimensionless quantities $\omega = f/\Delta f$ and $\xi = \alpha/\Delta\alpha$. While the initial prior for Model I was pessimistically chosen to be uniform in a large interval, when we restart the algorithm to test Model II, we can leverage upon the information already obtained for the parameters. In particular, we start with a precise estimation of the Rabi frequency $\bar{\omega} = \langle p(\omega) \rangle = 0.436$ with associated error $\bar{\sigma}_\omega$. The distribution of the chirping parameter is chosen to be centred in $\xi = 0$; a value for which Model II is equivalent to Model I. Although in general we may have no information about the variance of the new parameters added to improve a model, in this case it is possible to start with a rough estimate of an initial $\bar{\sigma}_\xi^2$. In fact, as Model I starts to provide predictions in contradiction with the experimental data when $\sigma_\omega \simeq \bar{\sigma}_\omega$, which corresponds to evolution times $t^* \simeq 1.26/\bar{\sigma}_\omega$ (using the PGH for QLE), for evolution times of this magnitude we can expect the frequency shift $t^*|\xi|$ to be non-negligible compared to the estimated Rabi frequency $\bar{\omega}$. This provides information on the interval range in which we can expect the correct value of ξ to be, that is $|\xi| \lesssim \bar{\omega}/t^* \simeq 10^{-8}$. Using this knowledge we set the initial prior probability distribution $p(\omega, \xi)$ to be the two-dimensional Gaussian distribution with mean $\mu = (\bar{\omega}, 0)$ and covariance

$$\Sigma(\omega, \xi) = \begin{pmatrix} (1.5 \times \sigma_\omega)^2 = 9.6 \times 10^{-5} & 0 \\ 0 & 10^{-8} \end{pmatrix}, \tag{7.2.9}$$

as shown in Fig. 7.10a, and restart the QLE algorithm to test Model II. For this two-parameters model the QLE algorithm was run offline using a classical simulator for the evolution with time-dependant Hamiltonians $\hat{H}(f, \alpha; t)$ and the experimental data from the NV$^-$ centre. The final distribution $p(\omega, \xi)$ obtained after 200 steps is reported in Fig. 7.10b. Interestingly, the evolution during the algorithm of the covariance norm $\|\Sigma\|_2$, shown in Fig. 7.9(inset), does not present a saturation within the 200 steps performed, and is able to achieve a precision higher than the final value $\|\Sigma\|_2 = 7.5 \times 10^{-6}$, which is almost an order of magnitude better than what achieved by Model I. Note that most probably, if we would continue to update the distribution with more experimental data, also for Model II we will meet a saturation level at some point, as even this model is very likely to not fully entail all the physics of the system under study. Nevertheless, a saturation level lower than the one for Model I is a first indication of the enhanced

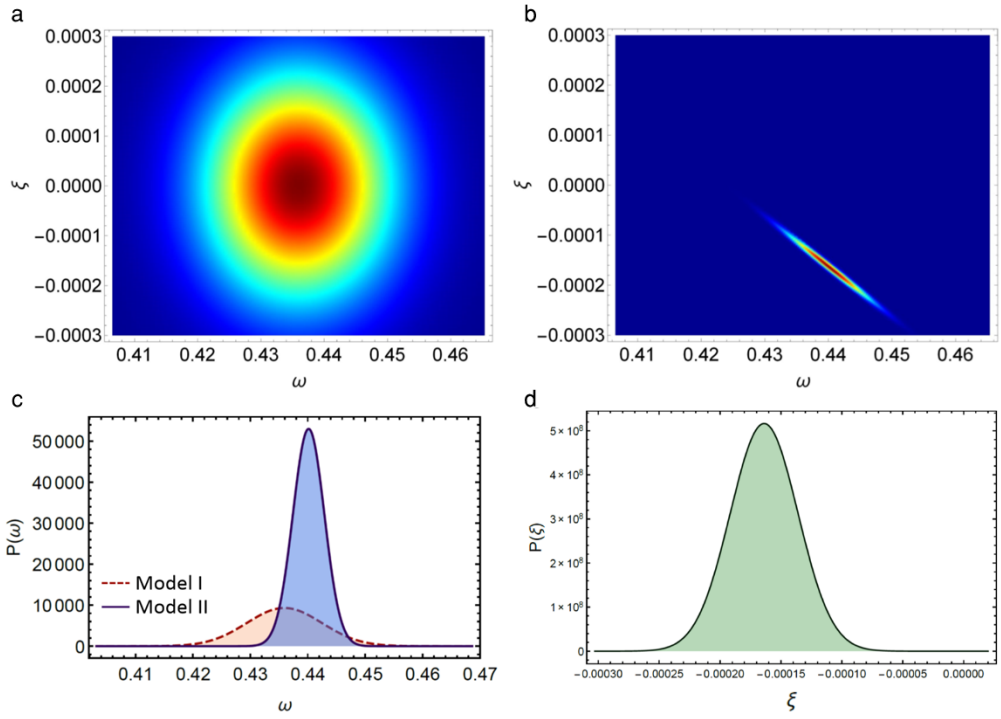


Figure 7.10: Learning and improving the model via QHL. **a,** Initial prior distribution of the two Hamiltonian parameters of the new model $\hat{H}(\omega, \xi)$ (Model II), which takes into account the information already obtained performing QLE on the initial model (Model I). **b,** Final probability distribution of the two Hamiltonian parameters after 200 steps of QLE. **c,** Final probability distribution for the single Rabi frequency parameter ω for Model II, compared with the one obtained for Model I. **d,** Final probability distribution for the chirping parameter ξ.

performance of Model II. The distributions of each single parameter, obtained from marginalising $p(\omega, \xi)$, are reported in Figs. 7.10c-d, and give $\omega_{\text{QLE}} = 0.440 \pm 0.002$ and $\xi_{\text{QLE}} = -1.6 \times 10^{-4} \pm 0.3 \times 10^{-4}$, corresponding to $f_{\text{QLE}} = (7.00 \pm 0.04)$ MHz and $\alpha_{\text{QLE}} = (-0.26 \pm 0.04)$ MHz^2. These values are consistent with the parameters $f_0 = (6.94 \pm 0.01)$ MHz and $\alpha_0 = (-0.28 \pm 0.03)$ MHz^2 calculated with a full fit of the Rabi oscillations, for evolution times long enough so that the chirping is manifested (see Fig 7.11).

Given the final distributions for both Model I and Model II, a formal comparison between them is given by the *Bayes factor* [18], defined as the ratio between the average likelihoods:

$$\mathcal{R} = \frac{\langle p(\mathcal{D}|f,\alpha)\rangle_{\text{Model II}}}{\langle p(\mathcal{D}|f)\rangle_{\text{Model I}}} = \frac{\int p(f,\alpha)p(\mathcal{D}|f,\alpha)df\,d\alpha}{\int p(f)p(\mathcal{D}|f)df}, \tag{7.2.10}$$

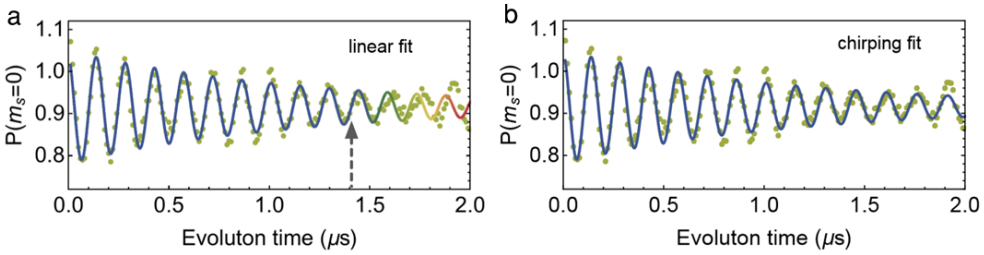


Figure 7.11: Rabi oscillations for longer evolution times. **a,** A linear fit with no time-dependant parameter is capable of accurately describe the observed dynamics for short times, obtaining a frequency $f_0 = (6.9 \pm 0.01)$ MHz. However, for evolution times approximately longer than 1.4 μs (dashed arrow) a discrepancy between the fit and the experiment data can observed, highlighted as coloured parts. **b,** More accurate predictions for longer evolution times can be obtained via a time-dependant model with $f'(t) = f + \alpha t$, where α is a chirping parameter. For this model, the best fit parameters are $f_0 = (6.94 \pm 0.01)$ MHz and $\alpha_0 = (-0.28 \pm 0.03)$ MHz^2.

where $\mathcal{D}$ represents the full set of data obtained from the NV$^-$ centre and $p(f)$ and $p(f, \alpha)$ are the final distributions for Model I and Model II respectively. Although the Bayes factor naturally favours simpler models, i.e. with fewer parameters, in our case we calculate a Bayes factor of $\mathcal{R} = 560$, which gives very strong evidence supporting Model II. This points out that a more complex model was necessary to provide a better description of the system. Therefore, our protocol for model learning concludes updating the model from Model I to Model II.

This was a very rudimentary and preliminary demonstration of a quantum model learning protocol, which is a topic of current research. Nonetheless, our demonstration already shows the basic underlying idea that QHL can be used not only to estimate the best model parameters, but also to instruct us on how to improve the model itself, providing potentially crucial insights into underpinning physical processes.

7.2.5 *Experimental results for IQLE*

To demonstrate IQLE, a crucial experimental difficulty is the implementation of a coherent quantum channel between the tested quantum system and the quantum simulator. In our experiment, this would have required the engineering of a quantum channel between the NV$^-$ centre and the photonic chip. While this could be achieved in principle, e.g. by exploiting spin-photon entanglement in the NV$^-$ centre (similarly to the approach in Ref. [19]) and using an optical link between the NV$^-$ centre and the photonic chip, it involved daunting challenges in our setup. For instance, the silicon material, where the simulator is embedded, does not transmit light at the visible wavelengths emitted by the NV$^-$ centre.

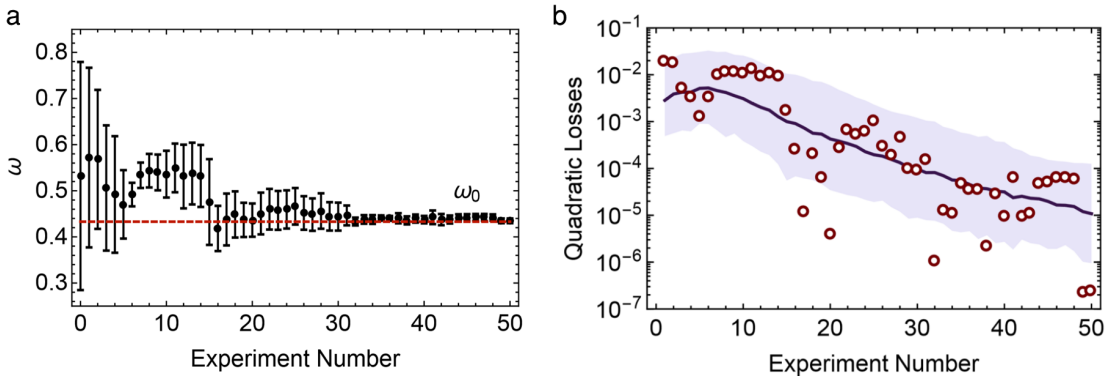


Figure 7.12: Experimental results for IQLE. **a,** Evolution of the mean and standard deviation of the distribution of the rescaled frequency ω, while IQLE is converging to the expected value of the ω_0 (dashed red line). Error bars are ± 1 s.d. of the distribution. **b,** Exponential decrease of quadratic losses for IQLE. Experimental data are shown as circles, and theoretical simulation data are shown as a line with a 67.5% confidence interval (shaded area). The theoretical simulation was averaged over 500 runs of IQLE.

For these reasons we instead opted to implement IQLE entirely on the photonic chip. In fact, the required quantum channel naturally can be simply implemented in a single integrated quantum photonics device using waveguides. In this case, IQLE can be applied to use calibrated gates to efficiently characterise other un-calibrated gates on the same chip, which now respectively represent the trusted hardware and the untrusted system to be validated. Therefore, we here exploit IQLE in the context of characterisation and verification of quantum devices. The operation to be characterised here is a $e^{-if_0 t\hat{\sigma_x}/2}$ evolution implemented on the $\hat{V}$ unitary of the chip (see Fig. 7.5f), where f_0 matches the value of the fitted Rabi frequency of the NV centre (chosen for consistency with the previous QLE demonstration). Characterising this $\hat{\sigma}_x$-rotation operation is thus equivalent to learning the Rabi frequency. To do it using IQLE, the inversion $e^{if_0 t\hat{\sigma_x}/2}$ was performed as the $\hat{U}$ unitary in the photonic chip, which allows us to obtain the output probabilities for IQLE experiments as described in section 7.2.2. The demonstration then proceeded similarly to QLE, but now using the chip also to perform experiments as the target system. That is, in each step of IQLE the experiment was implemented twice: once for measuring the outcomes from the untrusted $\hat{\sigma}_x$-rotation (top part in Fig. 7.3), and once for estimating the likelihoods (bottom part in Fig. 7.3).

Results for the estimated ω, as given by the posterior mean and standard deviation at each step of IQLE, are reported in Fig. 7.12a. The particle distribution converges quickly to the correct value ω_0. After 50 algorithm steps we obtain $f_{\text{IQLE}} = (6.92 \pm 0.08)$ MHz, which is within 1 s.d. of the implemented Rabi frequency $f_0 = 6.90$ MHz. The evolution of the quadratic losses (Fig. 7.12b) indicates that the parameter is learned exponentially

fast, with a final quadratic loss value of approximately 10^{-7}. The convergence of the algorithm to the implemented value ω_0 indicates the successful self-verification of the quantum device.

7.3 Discussion

In this chapter we have discussed Hamiltonian learning as a powerful tool for the characterisation and verification of quantum systems and devices. In contrast to other characterisation methods, e.g. standard tomographic approaches, quantum Hamiltonian learning exploits the use of quantum simulation to make the inference efficient, in an interesting combination of classical machine-learning methods and quantum technologies.

We also reported the first demonstration of QHL using quantum resources, showing its capability of validating Hamiltonian models and verifying quantum devices. While the experiments shown here use a digital quantum photonic simulator for the demonstration, QHL is universal and can be implemented on any quantum computing platform (e.g. photonics, trapped ions, superconducting and spin qubits), including analog simulators. This is particularly of interest because, as discussed in chapter 5, certain classes of analog quantum simulations are likely to approach a regime beyond that available to classical supercomputers earlier than digital approaches. With anticipated future developments in quantum hardware, the QHL protocol can be scaled up to learn more complex Hamiltonians, and promises the early delivery of quantum-enhanced computational techniques to efficiently characterise and verify quantum systems and technologies.

Finally we discussed a path to generalise QHL into a problem where instead of fixing the Hamiltonian model, we learn the underlying model itself. Although the results reported in this direction are very preliminary, the potential impact of algorithms able to tackle such problems is significant. It can be in fact argued that algorithms able to find new improved models to describe observed data could enable machines that can discover new physics in an automated manner. Although such tasks appear very challenging, we can now use QHL as a powerful tool to tackle it, as was preliminarily shown here for the simple case of a NV^- centre system. Studying this type of problem appears to be an interesting research direction, which is currently being explored.

Acknowledgements

The author's main contributions to the material reported in this chapter are the development of the code to implement the QHL algorithms, preliminary simulations and feasibility study of the experiment, design of the experiment, building the classical interface between NV^- centre and the photonic setup, implementation of the experiments in the photonic chip, implementation of the off-line algorithm for model learning, and data analysis. The implementation of the experiments in the NV^- centre setup was performed by Dr. Sebastian Knauer. The design of the experiments and the implementations on the photonic chip were done in cooperation with Dr. Jianwei Wang and Dr. Raffaele Santagati, with additional help from Mr. Antonio A. Gentile. Theoretical support was provided by Dr. Nathan Wiebe, while Dr. Maurangelo Petruzzella provided technical help in performing the preliminary simulations. The chip used for the experiments was fabricated by Toshiba. The experiment reported is published in Nature Physics [1].

References

1. Wang, J., Paesani, S., Santagati, R., Knauer, S., Gentile, A. A., Wiebe, N., Petruzzella, M., O'Brien, J. L., Rarity, J. G., Laing, A. & Thompson, M. G. Experimental quantum Hamiltonian learning. *Nature Physics* **13,** 551–555 (2017).
2. Gerritsma, R., Kirchmair, G., Zähringer, F., Solano, E, Blatt, R & Roos, C. Quantum simulation of the Dirac equation. *Nature* **463,** 68 (2010).
3. Popper, K. *Alles Leben ist Problemlösen* (Piper Verlag, Munich, 1994).
4. Aaronson, S. & Arkhipov, A. BosonSampling Is Far From Uniform. *Quantum Information and Computation* **14,** 1383–1423 (2014).
5. Mahadev, U. Classical verification of quantum computations. *arXiv:1804.01082* (2018).
6. Granade, C. E., Ferrie, C., Wiebe, N. & Cory, D. G. Robust online Hamiltonian learning. *New Journal of Physics* **14,** 103013 (2012).
7. Wiebe, N., Granade, C., Ferrie, C. & Cory, D. G. Hamiltonian Learning and Certification Using Quantum Resources. *Phys. Rev. Lett.* **112,** 190501. https://link.aps.org/doi/10.1103/PhysRevLett.112.190501 (2014).
8. Haake, F. *Quantum signatures of chaos* (Springer-Verlag, New York, 2004).
9. Santagati, R., Gentile, A. A., Knauer, S., Schmitt, S., Paesani, S., Granade, C., Wiebe, N., Osterkamp, C., McGuinness, L. P., Wang, J., Thompson, M. G., Rarity, J. G., Jelezko, F. & Laing, A. Magnetic-Field Learning Using a Single Electronic Spin in Diamond with One-Photon Readout at Room Temperature. *Phys. Rev. X* **9,** 021019. https://link.aps.org/doi/10.1103/PhysRevX.9.021019 (2 2019).

10. Wiebe, N., Granade, C., Ferrie, C. & Cory, D. Quantum Hamiltonian learning using imperfect quantum resources. *Phys. Rev. A* **89,** 042314. https://link.aps.org/doi/10.1103/PhysRevA.89.042314 (2014).

11. Wiebe, N., Granade, C. & Cory, D. G. Quantum bootstrapping via compressed quantum Hamiltonian learning. *New Journal of Physics* **17,** 022005 (2015).

12. Kantas, N., Doucet, A., Singh, S. S. & Maciejowski, J. M. An overview of sequential Monte Carlo methods for parameter estimation in general state-space models. *IFAC Proceedings Volumes* **42,** 774–785 (2009).

13. Granade, C. & Wiebe, N. Structured filtering. *New Journal of Physics* **19,** 083014 (2017).

14. Wiebe, N. & Granade, C. Efficient Bayesian Phase Estimation. *Phys. Rev. Lett.* **117,** 010503. https://link.aps.org/doi/10.1103/PhysRevLett.117.010503 (2016).

15. Jelezko, F., Gaebel, T, Popa, I, Gruber, A & Wrachtrup, J. Observation of coherent oscillations in a single electron spin. *Physical review letters* **92,** 076401 (2004).

16. Togan, E., Chu, Y., Trifonov, A., Jiang, L., Maze, J., Childress, L., Dutt, M. G., Sørensen, A. S., Hemmer, P., Zibrov, A. S., *et al.* Quantum entanglement between an optical photon and a solid-state spin qubit. *Nature* **466,** 730 (2010).

17. Doherty, M. W., Manson, N. B., Delaney, P., Jelezko, F., Wrachtrup, J. & Hollenberg, L. C. The nitrogen-vacancy colour centre in diamond. *Physics Reports* **528,** 1–45 (2013).

18. Barber, D. *Bayesian reasoning and machine learning* (Cambridge University Press, 2012).

19. Hensen, B., Bernien, H., Dréau, A. E., Reiserer, A., Kalb, N., Blok, M. S., Ruitenberg, J., Vermeulen, R. F., Schouten, R. N., Abellán, C., *et al.* Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres. *Nature* **526,** 682 (2015).

8 CONCLUSION

The results achieved in this thesis are obviously not an arrival point for integrated quantum photonics, and in particular silicon quantum photonics, but rather represent intermediate important steps to push forward the limits of the technology. We have to consider that when most of the projects reported here started, in 2015, the state-of-the-art in silicon quantum photonics was the on-chip generation and manipulation of two-photons in interferometers with $\approx$ 10 components [1], with the first fully on-chip two-photon interference reported only one year before that [2]. In retrospect, the results of chapters 3 and 4 were therefore hardly conceivable only four years ago, which shows the potential of the integrated photonics approach.

We may then wonder what can be achieved in the next four years from now. At the moment, we still seem to be far from the ultimate limits of the silicon photonics technology, and it is likely that integrated circuits with, at least, an order of magnitude more components will be available without too much effort. From the discussion in chapter 4, this should bring us to regimes where tens of photons can be generated and processed on single silicon chips. Hopefully, this could already provide some small computational speed-ups in the applications discussed in chapters 5, 6 and 7. Given that such applications, e.g. quantum chemistry simulations, have a strong scientific and industrial relevance, the achievement of any small efficiency advantage in this field would represent a technological milestone. It is important to remark that, however, we don't expect any major quantum advantage (so called *quantum computational supremacy*) in this pre-fault-tolerant regime [3].

A longer term scenario, if we want to target a regime far from being accessible with classical machines, must include the development and implementation of rudimentary error corrections techniques. For photonics, such error correction methods will need to tackle especially photon losses and limited source efficiencies. It is likely that to implement such schemes significant technological advancement will be required. In particular, we

can identify two main technological directions for integrated quantum photonics. The first is to progress towards ultra-pure integrated sources of heralded photons, while the second consists in developing low-loss high-speed integrated feed-forward schemes. It is likely that a strong involvement with the classical silicon photonics community, where an immense amount of engineering is dedicated to the technology development (motivated by classical applications e.g. in the telecommunication industry), will be required to achieve these goals. In the mean time, there is plenty of fun physics to do with current devices.

References

1. Silverstone, J. W., Santagati, R., Bonneau, D., Strain, M. J., Sorel, M., O'Brien, J. L. & Thompson, M. G. Qubit entanglement between ring-resonator photon-pair sources on a silicon chip. *Nature Communications* **6,** 7948 (2015).
2. Silverstone, J. W., Bonneau, D., Ohira, K., Suzuki, N., Yoshida, H., Iizuka, N., Ezaki, M., Natarajan, C. M., Tanner, M. G., Hadfield, R. H., *et al.* On-chip quantum interference between silicon photon-pair sources. *Nature Photonics* **8,** 104 (2014).
3. Neville, A., Sparrow, C., Clifford, R., Johnston, E., Birchall, P. M., Montanaro, A. & Laing, A. Classical boson sampling algorithms with superior performance to near-term experiments. *Nature Physics* **13,** 1153 (2017).

A BACKGROUND ON QUANTUM MECHANICS

This appendix contains an essential review of quantum mechanics, describing the basic principles and properties of quantum systems that are used throughout the thesis.

A.1 Classical statistics

Quantum mechanics is a theory that has its basis in the description of the statistical properties of physical objects, and is therefore inherently a statistical theory. Some notions of Bayesian statistics are also applied for the inference methods in chapters 6 and 7.

A.1.1 *Random variables and probability distributions*

When performing an experiment of any kind, we do not generally know a priori what the outcome will be. This uncertainty can arise for different reasons: we may have a limited control or observability on the experiment itself, or use an incomplete model to describe its complex behaviour, or there may be some intrinsic stochasticity in the system itself. To describe such uncertainty on the outcome, which we denote as a *random variable* $\mathtt{x}$, we associate to each of its possible values a *probability distribution* $p(\mathtt{x} = x) \equiv p(x)$, where $p(x)$ represents the probability of measuring the value x when performing the experiment. If $\mathcal{X}$ is the domain of the variable $\mathtt{x}$, i.e. the set of all possible values for the outcome x, then the probability distribution must satisfy $\sum_{x\in\mathcal{X}} p(x) = 1$.

Considering for simplicity $\mathcal{X}$ to be a discrete set, we can define the *expectation value*, or *mean*, of the the random variable $\mathtt{x}$ as $\langle x\rangle = \sum_{x\in\mathcal{X}} p(x)x$. More in general we can define the expectation value over functions $f(x)$ of the random variable as

$\langle f(x)\rangle = \sum_{x\in\mathcal{X}} p(x)f(x)$. In the case where $\mathcal{X}$ is a continuous set, rather than discrete, $p(x)$ is interpreted as a *probability density function*, and all the sums above are converted into integrals. An important quantity for probability distribution is the *variance* $\sigma^2(x) = \langle x^2\rangle - \langle x\rangle^2$. The square root of which is the *standard deviation*, which gives insights into how uncertain the outcome is. To generalise it to the case where outcomes are vector-valued $\mathbf{x}$, one can consider the covariance matrix $\Sigma(\mathbf{x}) = \langle \mathbf{x}\cdot\mathbf{x}^\top\rangle - \langle\mathbf{x}\rangle\cdot\langle\mathbf{x}^\top\rangle$, which describes the degree to which the variation in one component of the vector-valued random variable $\mathbf{x}$ depends on another component.

In the case where the experiment can provide more than one outcome simultaneously, for instance two outcomes $\mathtt{x}$ and $\mathtt{y}$, we can define a *joint probability distribution* $p(x,y)$ which denotes the probability that $\mathtt{x} = x$ and $\mathtt{y} = y$ simultaneously. Again, such probability distribution must satisfy $\sum_{xy} p(x,y) = 1$, where the sum is over all possible values of x and y. Given a probability distribution over a set of variables, we may be interested in knowing the probability distribution over just a subset of them. For example, given an experiment with two random variables $\mathtt{x}$ and $\mathtt{y}$ with joint probability distribution $p(x,y)$, we may be able to measure only the outcome of x without having access to y, and are thus only interested in $p(x)$. Such probability distribution for a subset of variables is called *marginal probability distribution*, and can be obtained my summing over the unmeasured variables

$$p(x) = \sum_y p(x,y), \tag{A.1.1}$$

an operation which is often called *marginalisation*.

A.1.2 *Conditional probability and Bayes' rule*

In many cases we are interested in understanding how one event affects another one. For example, the author may be interested in knowing what is the probability of finishing his PhD thesis in time given that he had a holiday to Venice. In this case we have two random variables $x \in \{\text{PhD}, \text{No PhD}\}$ and $y \in \{\text{Venice}, \text{No Venice}\}$ and we want to know what is the probability distribution for the variable $\mathtt{x}$ given that the event $\mathtt{y} = \text{Venice}$ has happened. This is called *conditional probability*, denoted via $p(\mathtt{x} = x|\mathtt{y} = y) \equiv p(x|y)$, which can be computed via

$$p(x|y) = \frac{p(x,y)}{p(y)}, \tag{A.1.2}$$

which is defined only for $p(y) \neq 0$. From the definition of conditional probability in eq. (A.1.2) and $p(x,y) = p(y,x)$ we can immediately arrive at Bayes' rule:

$$p(y|x) = \frac{p(x|y)p(y)}{p(x)} = \frac{p(x|y)p(y)}{\sum_y p(x|y)p(y)}, \tag{A.1.3}$$

where for the second equality we used $p(x) = \sum_y p(x,y) = \sum_y p(x|y)p(y)$ from the definition of conditional probability. Note that the denominator represents a normalisation coefficient, without which $p(y|x)$ would not be a probability distribution (we would have $\sum_y p(y|x) \neq 1$ in general).

Bayes's rule is a key element in studying inference processes as it essentially models our intuitive reasoning. Consider for example a child, Tom, playing Battleship. He has to decide whether to put his ships preferably in the centre of the board or towards the corners. His initial idea is that it is safer to not place them in the middle, and thus chooses to use the corners with higher probability $p(\text{corner}) = 60\%$ than for the second option, $p(\text{centre}) = 40\%$. However, after playing many rounds, Tom notices that, when using the corners, he manages to win only 2 every 10 rounds, i.e. $p(\text{win}|\text{corner}) = 20\%$. On the other hand, when he focuses on the centre of the board, he observes that half of the times he manages to win $p(\text{win}|\text{near}) = 50\%$. Based on this experience, applying Bayes' rule we can obtain the updated distribution $p(\text{centre}|\text{win}) = 62.5\%$ and $p(\text{corner}|\text{win}) = 37.5\%$. Therefore, in contrast to his initial belief, Tom has learnt that if he wants to win he would better place the ships towards the centre of the board.

This simple example shows how Bayes' rule allows us to update our knowledge based on the experience. Although the explanatory case described above is quite simple, the same principles can be used to perform learning tasks in much more complex situations. For instance, Bayesian inference methods underlie many modern machine-learning techniques, e.g. face recognition and cleanup of images [1]. Some of these methods are applied in the context of quantum algorithms in chapters 6 and 7.

Classical statistics thus provides powerful tools to model our intuitive reasoning when having to deal with our usual experience. However, as we will shortly see, it strikingly fails to describe what we can experience with microscopic physical systems, leading to the strangeness of quantum mechanics.

A.1.3 *Failure of classical statistics to describe microscopic systems*

Probably the most famous experiment which can be used to expose in a simple way the inadequacies of classical descriptions is the double-slit (or Young's) experiment. The schematic of the experiment is shown in Fig. A.1: an object is fired against a panel, where it can pass through one of two slits, which we label A and B. Finally, a screen, positioned at a sufficient distance from the slits, detects the arrival position x of the object. We can first think of performing the experiment having only one of the two slots open, and keeping the other slit closed, as is shown in Fig. A.1a. By repeating the experiment many

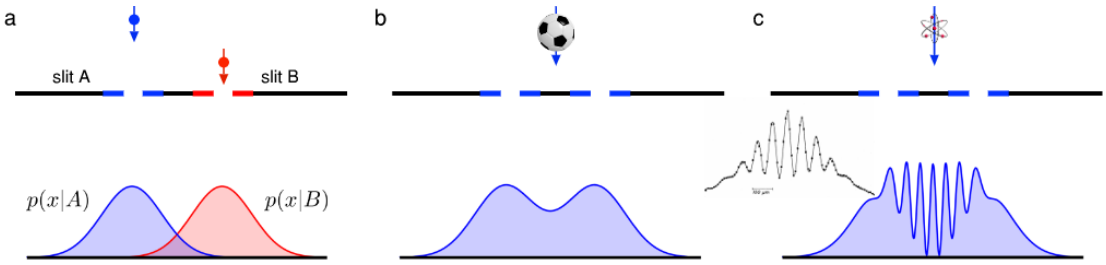


Figure A.1: Double-slit experiments. **a,** In the case the slit on which the particle passes through is known to be A (B), e.g. by keeping only the slit A (B) open, the probability distributions of the arrival point of the particle is $p(x|A)$ ($p(x|A)$). **b,** If the experiment is performed with classical objects, then the distribution observed with both slits open satisfies the classical statistical rule $p(x) = p(x|A)p(A) + p(x|B)p(B)$. **c,** If instead the experiment is performed with microscopic quantum objects, interference fringes are observed, and classical statistics fails to describe the output distribution $p(x) \neq p(x|A)p(A) + p(x|B)p(B)$. Inset: example of distribution experimentally observed in a double-slit experiment with neutrons (reported in Ref. [2]).

times we can measure the conditional distribution $p(x|A)$ of the final detected position in the case where only the slit A was opened (blue curve in Fig. A.1a), and similarly we can obtain $p(x|B)$ by opening B and closing A. These probabilities represent the distribution of the final position conditional on which slit the object has passed through. Now, what happens if we keep both slits opened? Classical statistics provides a simple answer to that. In fact, assuming that the object has equal probability to pass through A or B, i.e. $p(a) = p(b) = p_0$, using the rules for the conditional probability (A.1.2) and marginalisation (A.1.1), we simply have that the resulting distribution is proportional to the sum of the conditional ones: $p(x) = p_0[p(x|A) + p(x|B)]$, shown in Fig. A.1b. This is in fact what is observed when performing the experiment with conventional objects, e.g. firing footballs or bullets. However, when using microscopic physical objects, for example single electrons, the result changes radically. As pictured in Fig. A.1b, the distribution that is experimentally observed in a microscopic scenario presents fringes in the region where the two initial distributions overlap, indicating some form of interference. This behaviour is completely absent in the classical description, making the two resulting distributions significantly different[1]. In the last few decades such non-classical interference

1 In an attempt to recover a classical description of the fringes observed in the double-slit experiment, e.g performed with single electrons, one could try to explain it saying that the electron behaves like a classical wave, simultaneously passing through both slits, and thus generating classical interference fringes due to diffraction. However, this approach can be easily dismissed by inserting detectors right after the slits. When doing so in appropriate experimental conditions, it is observed that the detectors never click simultaneously, indicating that the electron actually passes only one of the two slots. Saying that the electron simultaneously passes through both slits is, therefore, erroneous: it

in the double-slit experiment has been demonstrated using systems at different scales, varying from single particles to complex massive molecules composed of up to hundreds of atoms [2–5].

Classical statistics thus fails to describe the observations that can be obtained from microscopic objects, which we can now start to call quantum systems. This manifests the need of a new theory to describe physical phenomena at the small scale. The next section is dedicated to recalling the new set of rules which can be used to describe such phenomena, forming the basics of the theory of quantum mechanics.

A.2 Basics of quantum mechanics

In this section we will recall some important definitions and properties of quantum states and measurement, and then proceed with the description of quantum dynamics.

A.2.1 *Quantum states*

Pure quantum states

In quantum mechanics, the physical state of a system corresponds to a normalised complex vector $|\psi\rangle$ in a Hilbert space $\mathcal{H}$. Two vectors are said to represent the same state if and only one is a multiple of the other, i.e. $|\psi_1\rangle = e^{i\phi}|\psi_2\rangle$ for some phase ϕ. Quantum states are thus defined up to a global phase. Any normalised vector in $\mathcal{H}$ corresponds to a valid physical state, which implies the possibility to superpose states $|\psi\rangle = \alpha|\psi_1\rangle + \beta|\psi_2\rangle$, where α and β, called *amplitudes*, are complex numbers. To satisfy the normalisation condition, we must have $|\alpha|^2 + |\beta|^2 = 1$. The presence of complex amplitudes in the quantum superposition is reminiscent of a wave behaviour, and is in fact what causes the observation of fringes in double-slit experiments.

The *dimension* of a quantum system is identified by the dimension of the Hilbert space associated to the system $d = |\mathcal{H}|$. Two-dimensional systems ($d = 2$) are called *qubits*, i.e. quantum bits, while general finite-dimensional systems with $d > 2$ are named *qudits*. In a d-dimensional Hilbert space we will usually define a reference basis as *computational basis*, and write it as $\{|0\rangle, |1\rangle, \ldots, |d-1\rangle\}$. Such basis is usually chosen as the most convenient reference for the experimental setup probing the system.

The state vectors described above are typically named as *pure* states, indicating a perfect knowledge of the system state.

cannot be described as a wave. The behaviour is somehow an hybrid between the classical concepts of a particle and a wave, representing the famous wave-particle duality.

Mixed states

When describing a quantum system, situations where an observer could not have a perfect knowledge of the state are also possible. For example, we can consider a situation where, possibly due to a limited control in an experiment, we have a probability p_1 to produce a state $|\psi_1\rangle$ and probability $p_2 = 1 - p_1$ for a different state $|\psi_2\rangle$. In this case the observer can describe the system in a *mixed* quantum state using the *density matrices* formalism.

A density matrix ρ is an operator on the Hilbert space $\mathcal{H}$ of the system which can be constructed as follows. If a system is in a pure state $|\psi\rangle$, its density matrix is defined as $\rho = |\psi\rangle\langle\psi|$. Given an ensemble of states $\{|\psi_i\rangle\}$, such that the system has probability p_i to be in $|\psi_i\rangle$, the mixed state of the system is given by the density matrix $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$. It is straightforward to verify that such matrix is Hermitian ($\rho^\dagger = \rho$), positive semi-definite ($\rho \geq 0$), and trace-1 ($\operatorname{Tr}\rho = 1$). Any density matrix satisfying these three properties represents a valid physical mixed quantum state.

A quantity useful to quantify the mixedness of a quantum state is the *Von Neumann entropy*, defined as $\mathcal{S}(\rho) = -\operatorname{tr}(\rho \ln \rho)$. In the case where, as above, the density matrix is given by a probability distribution p_i over an ensamble of states $\{|\psi_i\rangle\}$, and if this ensamble forms an orthogonal set of states, then it is easy to see that the Von Neumann entropy reduces to $\mathcal{S}(\rho) = -\sum_i p_i \ln p_i$, which is the Shannon entropy of the distribution p_i. In particular, it is zero if the state is pure, i.e. all the p_i are zero except one, and is maximum when the p_i distribution is uniform. A simplified for of the Von Neumann entropy is its linearised form, where we substitute $\ln\rho \to \mathbb{1} - \rho$ and obtain, up to a constant, the so called *purity* of a quantum state $\mathcal{P}(\rho) = \operatorname{tr}\rho^2$. It can again be easily observed that the purity is one if and only if the state is pure, and has its minimum value $1/d$ if its given by a uniform superposition of an orthogonal basis. In the latter case, where $\rho = \mathbb{1}/d$, we say that the state is maximally mixed.

Geometrical representation of qubit states: the Bloch sphere

Quantum states associated to a qubit system admit a natural geometrical representation. The density operator of a single qubit is a 2×2 matrix. As the three Pauli matrices and the identity form a basis for the 2×2 matrices, we can always write $\rho = a\mathbb{1} + b\sigma_x + c\sigma_y + d\sigma_z$. By the linearity of the trace, and using that the Pauli matrices have null trace, the condition $\operatorname{tr}\rho = 1$ gives $a = 1/2$. We can then write

$$\rho = \frac{\mathbb{1} + \boldsymbol{r}\cdot\boldsymbol{\sigma}}{2}, \tag{A.2.1}$$

where $\boldsymbol{r} = (x, y, z)$ is a real three-dimensional vector which uniquely define the state, called the *Bloch vector*, and $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$. The positivity and condition implies that

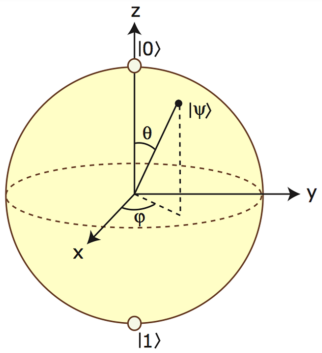


Figure A.2: Bloch sphere representation of a qubit state. Vectors on the sphere surface represent pure quantum states, while vectors contained in the sphere are mixed states. The centre is associated to the maximally mixed state of a qubit.

$|\boldsymbol{r}| \leq 1$, providing a nice geometrical representation of the state of a qubit. Furthermore, one can easily calculate that the norm of the Bloch vector is directly related to the purity of the state via $\mathcal{P} = (1 + |\boldsymbol{r}|^2)/2$, which means that the states on the surface of the sphere $|\boldsymbol{r}| = 1$ represents all and only pure states. The mixedness increases towards the centre of the sphere, with $|\boldsymbol{r}| = 0$ representing the maximally mixed state.

Fidelity of quantum states

Having defined quantum states, it is helpful to have a quantity to measure "how close" two quantum states ρ and σ are, i.e. ideally a distance between quantum states. A good measure is the quantum state *fidelity*, defined as

$$F(\rho, \sigma) = \left(\mathrm{Tr} \sqrt{\rho^{1/2} \sigma \rho^{1/2}} \right)^2 . \tag{A.2.2}$$

Although it is not evident in this definition, the fidelity is symmetric in the input states ρ and σ [6]. If one of the two states is pure, e.g. $\sigma = |\psi\rangle \langle\psi|$, the fidelity reduces to

$$\begin{aligned} F(|\psi\rangle, \rho) &= \left(\mathrm{Tr} \sqrt{\langle\psi| \rho |\psi\rangle |\psi\rangle \langle\psi|} \right)^2 \\ &= \langle\psi| \rho |\psi\rangle . \end{aligned}$$

Finally, if both states are pure we can also write $\rho = |\phi\rangle \langle\phi|$ for some state $|\phi\rangle$, and obtain

$$F(|\psi\rangle, |\phi\rangle) = |\langle\psi|\phi\rangle|^2. \tag{A.2.3}$$

which is simply the projection of one state on the other and represents, as we will shortly describe, the probability of measuring a system, prepared in $|\psi\rangle$, in the state $|\phi\rangle$. In this form the fidelity is clearly symmetric, and presents some features we expect from

a measure of distance. In particular, $F(|\psi\rangle, |\phi\rangle) = 1$ if and only if $|\psi\rangle = |\phi\rangle$, and $F(|\psi\rangle, |\phi\rangle) = 0$ if $|\psi\rangle$ and $|\phi\rangle$ are orthogonal.

States of composite quantum systems

A composite system of two of or more subsystems is described via a *tensor product* structure of the the individual Hilbert spaces. For two subsystems with individual Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$, the composite states $|\psi_{AB}\rangle$ will thus live in the Hilbert space $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$. If the two systems are individually prepared in the pure states $|\psi_A\rangle$ and $|\psi_B\rangle$, then the global state is given by $|\psi_{AB}\rangle = |\psi_A\rangle \otimes |\psi_B\rangle$. When considering a system composite of N subsystems, each with dimension d, the dimension of the total system $\mathcal{H}$ is d^N. Mixed states are again described by the density matrix formalism.

We may be interested in describing only the state of a subset of subsystems, independently from the state of the remaining ones. Analogously to the marginalisation procedure in classical statistics, in quantum mechanics the procedure is to "trace out" the remaining subsystems. That is, considering again a bipartite system, the density matrix of the individual subsystem A is given by:

$$\rho_A = \mathrm{Tr}_B(\rho_{AB}), \tag{A.2.4}$$

where ρ_{AB} is the density matrix of the total bipartite system, and Tr_B is the *partial trace* over system B, defined as $\mathrm{Tr}_B(\rho_{AB}) = \sum_i {}_B\langle\phi_i|\rho_{AB}|\phi_i\rangle_B$ with $\{|\phi_i\}_B$ an orthonormal basis of the Hilbert space $\mathcal{H}_B$.

A.2.2 *Measurements on quantum systems*

Measurement postulates and Born's rule

The statistics of a measurements performed on a quantum state are dictated by the *Born's rule*, which can be considered as an axiom of quantum mechanics. It states that, if we prepare a system in a state $|\psi\rangle$, then the probability of observing the system in a state $|\psi\rangle$ when measuring it is given by $p(|\phi\rangle)_{|\psi\rangle} = |\langle\phi|\psi\rangle|^2$. This probability can be considered as a classical probability of measuring the outcome $|\phi\rangle$ conditional on the initial preparation of the state $|\psi\rangle$. Another important aspect of measurement of quantum mechanics is that measurements on a system disturb its state. In fact, if the system is initially in the state $|\psi\rangle$, upon the detection of $|\phi\rangle$ the state of the system is not longer $|\psi\rangle$, but is instantaneously updated to $|\psi'\rangle = |\phi\rangle$. This instantaneous update of the state is the well-known *collapse of the wave-function*.

Observables and projective measurements

In quantum mechanics an observable is represented by an Hermitian linear operator on the Hilbert space of the system, and the outcome of the observation is one of its eigenvalues. Due to the spectral theorem, requiring the observable to be Hermitian ensures that its always diagonalisable via an orthonormal basis, i.e. a basis of physical states, with real-valued eigenvalues. This makes the theory consistent with our experience, where no complex outcomes are obtained from experiments. Let us then consider an observable A with real eigenvalues $\{a_k\}$ and relative orthonormal basis of eigenvectors $\{|k\rangle\}$. We can then write A as $A = \sum_k a_k |k\rangle \langle k| = \sum_k a_k \Pi_k$, were $\Pi_k = |k\rangle \langle k|$ are projective operators on the eigenvectors. Born's rule allows us to calculate the statistics of the outcomes. In fact, if the system is in the state $|\psi\rangle$, the probability of obtaining a_k as the outcome of the measurement is $p(a_k) = |\langle k|\psi\rangle|^2 = \langle\psi| \Pi_k |\psi\rangle$. More in general, if the system is prepared in a mixed state ρ we have $p(a_k) = \langle k| \rho |k\rangle = \mathrm{Tr}(\Pi_k \rho)$. The expectation value for the observable is then simply given as the average over all possible outcomes $\langle A\rangle = \sum_k p(a_k) a_k = \langle\psi| A |\psi\rangle = \mathrm{Tr}(A\rho)$.

To include the collapse of the wave-function, if a measurement provides an outcome $a_{\tilde{k}}$, the state of the system is immediately updated to the associated eigenstate $|\tilde{k}\rangle$. More formally, the state instantaneously becomes

$$|\psi'\rangle = \frac{\Pi_{\tilde{k}} |\psi\rangle}{\sqrt{\langle\psi| \Pi_{\tilde{k}} |\psi\rangle}}, \tag{A.2.5}$$

where the denominator is required to maintain the normalisation of the state. In general we can also think of not projecting on a full basis, but on subspaces of the Hilbert space. For example, in chapter 7 we will be interested in projecting into a vector $|\psi\rangle$ and on its orthogonal subspace $|\psi\rangle^\perp$. Both situations described above are part of the *projective measurement* class, which represents the type of measurements described by a set of projective operators $\{\Pi_k\}$ that obeys $\sum_k \Pi_k = \mathbb{1}$. This last property ensures that the probabilities $p(a_k) = \mathrm{Tr}(\Pi_k \rho)$ for the outcomes sum to one.

A.2.3 *Correlations in quantum states*

We have now seen the fundamental properties of quantum states and measurements. Some if the features encountered, such as the possibility to superpose states and change the state by measuring it, are counter-intuitive for our typically experience. However, we may still wonder if there is something fundamentally different in the quantum description, or if we could somehow come up with some complex classical theory that encompasses all these odd properties of quantum systems. Surprisingly, we will see in this section how any attempt to describe quantum mechanics via a classical structure is doomed to

failure. Such divergence between classical and quantum physics manifests itself in the correlations that can arise between quantum systems.

Quantum entanglement

The *entanglement* of quantum systems is probably the most eminent exotic aspect of quantum mechanics. Firstly recognised in a series of famous articles by Einstein and coworkers [7], and Schrödinger [8], it has been largely studied, reviewed and discussed in the last few decades [9, 10]. It is often regarded as the origin of the quantum enhancement in many applications [6].

To describe entanglement, for simplicity we consider bipartite systems. All the definitions that we will see can be easily extended for any multi-partite system, although interpretations get quite more intricate. Let us thus consider two subsystems A and B with Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$. An arbitrary bipartite pure state on the joint space can be written as $|\psi\rangle_{AB} = \sum_{i,j} a_{ij} |i\rangle_A \otimes |j\rangle_B$, where $\{|i\rangle_A\}$ and $\{|j\rangle_B\}$ are bases of states for $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively (and therefore $\{|i\rangle_A \otimes |j\rangle_B\}$ is a basis of the joint space), and $\sum_{ij} |a_{ij}|^2 = 1$.

The separability and the entanglement for a pure state are defined as follows. A pure state $|\psi\rangle_{AB}$ is *separable* if there exist two vectors $|\Lambda\rangle_A \in \mathcal{H}_A$ and $|\chi\rangle_B \in \mathcal{H}_B$ such that $|\psi\rangle_{AB} = |\Lambda\rangle_A \otimes |\chi\rangle_B$, otherwise it is *entangled*. Examples of pure entangled states for a bipartite system of qubits are the four Bell states

$$|\Phi^{\pm}\rangle_{AB} = (|0\rangle_A |0\rangle_B \pm |1\rangle_A |1\rangle_B)/\sqrt{2} \tag{A.2.6}$$

$$|\Psi^{\pm}\rangle_{AB} = (|0\rangle_A |1\rangle_B \pm |1\rangle_A |0\rangle_B)/\sqrt{2}. \tag{A.2.7}$$

It is in fact straightforward to notice that there is no way these states can be separated into two states of the individual subsystems, and are hence entangled.

The definition of entanglement can be extended for mixed states ρ_{AB} saying that a mixed state ρ_{AB} is separable if it can be written as

$$\rho_{AB} = \sum_i p_i \rho_i^{(A)} \otimes \rho_i^{(B)} \tag{A.2.8}$$

for some sets $\{\rho_i^{(A)}\}$ and $\{\rho_i^{(B)}\}$. Otherwise it is entangled. An example of an entangled mixed state of two qubits is the Werner state [11]

$$\rho_W = \frac{1-p}{4} \mathbb{1} + p \, |\Psi^-\rangle \langle \Psi^-| \,, \tag{A.2.9}$$

given by a mixture of the completely mixed state $\mathbb{1}/4$ and the Bell state $|\Psi^-\rangle$. It can be shown that this state is entangled for $\frac{1}{3} < p \leq 1$ [11].

The meaning of having entanglement indicates, roughly speaking, the presence of non-classical correlations: correlations stronger than what could be possible to get using classical resources. To have a better intuition of this we can consider the following example. Consider now a bipartite separable state $\rho_{AB} = \rho_A \otimes \rho_B$. To describe each individual subsystem separately we can trace out the other part and trivially obtain:

$$\rho'_A = \mathrm{Tr}_B(\rho_{AB}) = \rho_A \,\mathrm{Tr}(\rho_B) = \rho_A, \tag{A.2.10}$$
$$\rho'_B = \mathrm{Tr}_A(\rho_{AB}) = \mathrm{Tr}(\rho_A)\rho_B = \rho_B. \tag{A.2.11}$$

This means that for a separable system it is absolutely equivalent to describe the single subsystem and to describe the total system as a whole. In other words, the information contained by the total system is equal to the information of the subsystems taken alone. This makes perfect sense with our experience, and is in fact true in classical physics: by analysing the each small part of a large system we can get all possible information. This is strikingly not true in presence of quantum entanglement.

Let us consider for example the Bell states $\left|\Phi^+\right\rangle$. In this case the states of the individual subsystems are

$$\rho'_A = \mathrm{Tr}_B\left(\left|\Phi^+\right\rangle\!\left\langle\Phi^+\right|\right) = \mathrm{Tr}_B\left[\frac{\mathbb{1}_A}{4}\left|0\right\rangle\left\langle 0\right|_B + \frac{\mathbb{1}_A}{4}\left|1\right\rangle\left\langle 1\right|_B\right] = \frac{\mathbb{1}_A}{2}, \tag{A.2.12}$$

and similarly $\rho'_B = \frac{\mathbb{1}_B}{2}$. Then, even if the composite system is in a pure state, i.e. we have a all possible information about it, the states of the subsystems are completely mixed, i.e. no information is available. This means that, by describing the two subsystems separately, we would miss a lot of information about the total entangled state. This is due to the presence of non-classical correlations between A and B, correlations with no classical analog, which we will shortly analyse in more details.

From the example above it thus appears that entanglement in a pure state induces uncertainty in the states of the subsystems. This fact can be used to quantify entanglement via the mixedness of a subsystem, which is in turn measured via the Von Neumann entropy $\mathcal{S}$ (see A.2.1). Such quantifier of entanglement is called *entropy of entanglement*, defined by $\mathcal{E}(\rho_{AB}) = \mathcal{S}(\rho_A) = \mathcal{S}(\rho_B) = \mathcal{S}(\mathrm{tr}_A\left(\rho_{AB}\right))$. To make the definition consistent, we have used a symmetry property of the Von Neumann entropy, which is independent from which subsystem the partial trace is performed on $\mathcal{S}(\mathrm{tr}_B\left(\rho_{AB}\right)) = \mathcal{S}(\mathrm{tr}_A\left(\rho_{AB}\right))$ [6]. The states which possess higher entanglement, as quantified by the entropy of entanglement, are those who provide partial density matrix that are maximally mixes. Such states are called *maximally entangled*. As we have seen in the example before, the Bell states are maximally entangled states of two qubits.

Apart from the entropy of entanglement, a wide variety of entanglement quantifiers have been proposed. A comprehensive review of them can be found in Ref. [12].

Bell non-locality

As we have just seen, the presence of quantum entanglement indicates the presence of correlations that do not intuitively exist in classical systems. But what are these correlations, and what do we exactly mean by calling them non-classical? Related to this question is the fact that, while we have introduced quantum states and measurements to describe quantum mechanical systems, we still have not ruled out the possibility that a more complex classical theory could be proposed to describe quantum mechanical properties. While such possibility may sound unreasonable, it was considered very seriously in the early days of quantum mechanics, and, for example, was famously championed by Eistein, Podolski and Rosen in their celebrated paper [7]. It was only in 1964 that Bell proposed an experiment test that can unequivocally rule out any possible classical explanation of Nature, and definitely prove that quantum mechanics is fundamentally different from any classical formulation [13]. Such type of tests are now called a *Bell inequalities*.

An important feature about Bell inequalities is that they are not result about quantum mechanics, but rather on the capabilities of generic classical theories to describe the correlations that can happen in Nature. They are therefore independent of the formalism of quantum mechanics, and are thus derived disregarding the description of quantum states and measurements that we have discussed so far. Let us consider a scenario where two separate parties, Alice and Bob, can perform experiments on their respective set-ups. The experiment for Alice consists in choosing randomly which operator to measure one between two options x_1 and x_2, and then collect the outcome $a \in \{a_0, a_1, \ldots\}$ of her measurement. For example, for these observables Alice could decide x_1 to be the position of a particle, and x_2 the momentum, or x_1 as the spin along the $\hat{z}$ direction and x_2 along the $\hat{y}$ direction. Similarly, Bob can choose between two observables y_1 and y_2, and observes the outcome $b \in \{b_0, b_1, \ldots\}$. Such scenario is pictured in Fig. 3.6. The laboratories of Alice and Bob are considered to be independent, for example by being separated by many light-years, so that both the choice of the measurements in one laboratory does not depend on what happens in the other. By repeating the experiments many times they can record the probability distribution for the measurement outcomes $p(ab|xy)$, with $x \in x_1, x_2$ and $y \in y_1, y_2$.

Now, how would we write the probability distribution $p(ab|xy)$ in a general classical description? In classical physics, given that the two laboratories are independent, the outcome in one experiment does not depend on the measurement performed in the other. This is a perfectly reasonable assumption intuitively, which is completely consistent with our everyday experience. Under this assumption we can factorise the probability distribution as

$$p(ab|xy) = p(a|x)p(b|y). \tag{A.2.13}$$

Any model which allows us to factorise in this way the joint distribution is called *local theory*. As a further step, to make the classical description as general as possible, we have to include the possibility that Alice and Bob may have met in space at some point in their history, which could have induced some shared correlations between the two systems. These shared correlations can be encoded into an auxiliary variable λ, called *hidden variable*, which accounts for any hidden pre-programming in the two systems. Using these two assumptions we have a *local hidden-variable theory*, where the general form of the probability distribution is

$$p(ab|xy) = \int_\Lambda p(a|x,\lambda)p(b|y,\lambda)p(\lambda)\ d\lambda, \tag{A.2.14}$$

where we also consider the variable λ to be governed by a probability distribution $p(\lambda)$ over its support Λ. Note that, to arrive to eq. (A.2.14) we didn't make any particular choice of physical system, but only minimal assumptions on the fundamental properties of the underlying theory. Bell inequalities can be used to prove that the correlations that can be observed in quantum systems are incompatible with the eq. (A.2.14), and therefore that that classical local models cannot fully describe Nature. To see this in a simple case we can consider dichotomic measurements for Alice and Bob, where $a, b \in \{-1, +1\}$. In this configuration, Clauser, Horne, Shimony and Holt have shown that, if eq. (A.2.14) holds, than the following inequality (called *CHSH inequality* after its authors [14]) holds:

$$S = |\langle a_1 b_1\rangle + \langle a_1 b_2\rangle + \langle a_2 b_1\rangle - \langle a_2 b_2\rangle| \leq 2, \tag{A.2.15}$$

where the quantities $\langle a_\alpha b_\beta\rangle = \sum_{a,b} ab\ p(ab|x_\alpha y_\beta)$, called *correlators*, are the expectation values of the product of the outcomes for different choices of the measurement bases. Any local hidden-variable theory must satisfy this condition.

On the other hand, the predictions of quantum mechanics are not compatible with it. To see this, let us describe the Bell experiment scenario using the formalism for quantum systems discussed in the previous sections. In particular, Alice and Bob can share an entangled pair of qubits, for example the maximally entangled state $|\Psi^-\rangle = (|0\rangle_A |1\rangle_B - |1\rangle_A |0\rangle_B)/\sqrt{2}$. Such state could have been prepared from a third party, Charlie, who then sends the individual qubits to Alice and Bob, who then randomly choose which measurement to performed between the following operators:

$$x_1 = \sigma_z, \qquad y_1 = -\frac{\sigma_z + \sigma_x}{\sqrt{2}}, \tag{A.2.16}$$

$$x_2 = \sigma_x, \qquad y_2 = \frac{\sigma_z - \sigma_x}{\sqrt{2}}. \tag{A.2.17}$$

These are all Hermitian operators, and thus represent valid observables. The correlators can be readily calculated via $\langle a_\alpha b_\beta \rangle = \langle \Psi^- | (x_\alpha \otimes y_\beta) | \Psi^- \rangle$, which gives $\langle a_1 b_1 \rangle = \langle a_1 b_2 \rangle = \langle a_2 b_1 \rangle = -\langle a_2 b_2 \rangle = 1/\sqrt{2}$. The Bell-CHSH inequality is thus violated:

$$S = 2\sqrt{2} \geq 2. \tag{A.2.18}$$

This allows us to formally claim that quantum mechanics is fundamentally incompatible with all possible classical local hidden-variable theories. Experimental violation of Bell inequalities have been reported many times in the last few decades, starting from the 1980s [15], in a number of different experimental platforms. With the development of more sophisticated quantum technologies, experiments in conditions nearly identical to the ideal Bell's experiment have been achieved [16–18], closing the door to any reasonable argument for a classical description of Nature.

It is important to remark that the violation of a Bell inequality is not a result concerning quantum mechanics itself, but rather it is about the impossibility to describe correlations that can be observed in Nature using classical theories. Nonetheless, it shows that entangled quantum system can possess correlation such that the it is not possible to separate the join probability distribution as in (A.2.14). These correlations are called *non-local correlations*, and any system that do not satisfy (A.2.14) is said to be *Bell non-local*. The values of $S = 2\sqrt{2}$ can be shown to be the maximum violation of the CHSH inequality for bipartite states of qubits, known as the *Tsirelson's bound* [19]. For the CHSH inequalities, it can be achieved only using maximally entangled state. In fact, in the framework of quantum mechanics, non-locality is related to entanglement in the sense that it can be easily seen that any non-local quantum system must be entangled. The opposite is however not true in general: systems can be entangled without being non-local. An example is the Werner state (see eq. (A.2.9)), which for $\frac{1}{3} \leq p < 1/\sqrt{2}$ shows entanglement but does not violate Bell inequalities. Non-locality thus represent a stronger form of non-classical correlation compared to entanglement. Only for pure states the two concepts are equivalent [20].

In the discussion above we considered a bipartite Bell experiment scenario with two-outcome measurements. This can be generalised to multi-partite scenario and measurements with an arbitrary number of outcomes. A review of Bell inequalities different scenarios can be found in Ref. [20]. For example, in chapter 3 we have reported experimental violations of Bell inequalities for measurements with multiple outcomes using entangled qudits.

A.2.4 *Evolution of quantum system*

So far we have discussed static quantum systems, i.e. the formalism and the properties of quantum states and measurements. This section instead focuses the basic formalism to describe quantum dynamics.

Schrödinger's equation and unitary evolutions

For simplicity, we focus here on the evolution for closed quantum systems, i.e. systems that do not interact with any external one. In this case, the dynamics of a system can described via the Hamiltonian operator H. Considering a pure quantum system, the evolution of its state is given by the famous *Schrödinger's equation*:

$$i\hbar\frac{\partial}{\partial t}|\psi\rangle = H|\psi\rangle. \tag{A.2.19}$$

The validity of this differential equation is generally considered a postulate of the quantum theory. For systems in a mixed state, Schrödinger's equation can be rewritten as:

$$i\hbar\frac{\partial}{\partial t}\rho = -[\rho, H], \tag{A.2.20}$$

called the *Von Neumann's equation*. The general solutions of eq. (A.2.19) and eq. (A.2.20) can be written as

$$|\psi(t)\rangle = U(t)|\psi\rangle, \tag{A.2.21}$$

$$\rho(t) = U(t)\rho_0 U^\dagger(t) \tag{A.2.22}$$

with $U(t)$ a unitary operator. If the Hamiltonian is time-independent, the unitary evolution has the explicit form

$$U(t) = \exp(-iHt/\hbar). \tag{A.2.23}$$

In all these equations the quantity $\hbar$ is Planck's constant. For simplicity, the units of measure can be redefined so that $\hbar = 1$ can be used, as will often be done in this thesis.

Unitary operations are thus the tool to describe the evolution of quantum states. Note that, the only operation that is not unitary is the collapse of the quantum state when a measurement is performed. In fact, as we have seen in the previous section, those operations are described by projections; highly non-unitary operations. This opens an inconsistency in the theory, where measurements are treated differently from any other types of operations. This issue, know as the measurement problem, remains an open problem.

In the context of quantum computing, unitary operations are referred to as *gates*. Because any unitary operation on a quantum system of arbitrary dimension can be decomposed into single-qubit and two-qubit gates [6], the ability implement these two classes of quantum gates is sufficient to perform universal quantum computation.

Single-qubit gates

Classically the only non-trivial operation on a single bit(apart from the identity) is the `NOT` gate, which switches the value of a single bit. On the other hand, for a single qubit we can choose over a large variety of operations: any 2×2 unitary matrix can be used. The most common single qubit gates are the Pauli matrices σ_x, σ_y and σ_z (sometimes we may also simply refer to them as X, Y and Z), the Hadamard gate

$$H = \frac{1}{\sqrt{2}}(Z + X) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \tag{A.2.24}$$

and the phase shift operation

$$R_Z(\phi) = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\phi} \end{pmatrix}, \tag{A.2.25}$$

which represents a rotation by an angle ϕ along the $\hat{z}$ axis of the Bloch sphere. A rotation of an angle θ around a generic axis $\hat{n}$ of the Bloch sphere is given by the operator

$$U = e^{i\frac{\theta}{2}\hat{n}\cdot\sigma}. \tag{A.2.26}$$

In a multi-partite system with Hilbert space $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \cdots \otimes \mathcal{H}_N$, if we operate single qubit gates separately on each subsystem, the total operation is given by the unitary

$$U = U^{(1)} \otimes U^{(2)} \otimes \cdots \otimes U^{(N)}. \tag{A.2.27}$$

It is easy to notice from eq. (A.2.27) that, if we apply U to a separable state, it remains separable. Similarly, if the state is entangled we can't make it separable via local operations: operating single qubit gates on the subsystems of a multi-partite state,we cannot generate nor destroy entanglement. In other words, single qubit gates perform no interactions between the subsystems.

Two-qubit gates

In order to generate interactions, and hence manipulate entanglement, we to operate gates that cannot be written as single qubit gates acting separately as in eq. (A.2.27).

The simplest of such operations are two-qubit gates. A common type of two-qubit gates are the *controlled gates*: linear operations where a unitary operation U is performed on a target qubit (T) conditional on the state of a control qubit (C) being $|1\rangle$. That is, if the qubit C is in $|1\rangle_C$ then U is applied to T, if C is in $|0\rangle_C$ then the state of T is left unvaried. This operation represents a general controlled-unitary (`C-U`) gate, and its operator can be explicitly written as

$$\texttt{C-U} = |0\rangle\langle 0|_C \otimes \mathbb{1}_T + |1\rangle\langle 1|_C \otimes U_T. \tag{A.2.28}$$

Examples of controlled-gates are the `C-NOT` gate (where $U = X$) and the `C-Phase` gate (where $U = R_Z(\phi)$). To see that two-qubit gates can be used to generate entanglement, it is sufficient to consider an initial state $|+\rangle|0\rangle = (|0\rangle + |1\rangle)|0\rangle/\sqrt{2}$, which is clearly separable, and apply a `C-NOT` gate to it. The output state obtained is $\texttt{C-NOT}|+\rangle|0\rangle = (|00\rangle + |11\rangle)/\sqrt{2} = |\Phi^+\rangle$, that is a Bell state. We have hence generated a maximally entangled state starting from a product state.

Universal operations on arbitrarily many qubits can be performed combining elements from small, discrete sets of elementary operations on few qubits, called *universal gate sets* [6]. An example of such universal set is given by one two-qubit operation and two single-qubit operations: the `C-NOT`, the Hadamard gate, and the rotation $R_Z(\pi/4)$ (called `T` gate). All such gate sets to perform universal quantum computations must contain at least one entangling gate.

A.3 Quantum technologies

In the previous sections we have described that, when controlling individual with quantum mechanical systems, we can exploit rules that are quite different from what we is available with classical systems. As for any game, when the set of rules is changed, new opportunities arise. A major research topic in the last few decades has been to explore such opportunities to develop new quantum technologies: technologies based on quantum systems. The driving force behind it is the progress in the last few decades in engineering hardware able to control single quantum objects with increasing reliability and precision, which is now possible in a number of different platforms [21].

Such engineering progress promises to have applications in a wide range of technologies: from secure communications with secrecy ensured by the laws of quantum mechanics, to non-destructive high-resolution sensors based on single atoms. At the time of writing, the two quantum technologies just mentioned, *quantum cryptography* and *quantum sensing*, have already reached the market. Probably the most celebrated technology, but also the most daunting in terms of hardware requirements, is quantum computing. The excitement over quantum computers is due to the fact that they are strongly believed

to be able to solve important problems untackable with classical machines, such as the factorisation of large numbers and the simulation of complex chemical systems. The scope of this section is to discuss the basic concepts of quantum computers, setting a generic background to what have been presented in the rest of the thesis in the specific context of photonic quantum information processing.

A.3.1 *Quantum computing*

A computer can in general be described as a physical machine able to encode some well-defined inputs from an external user, process them, and produce some outputs readable by the user. The main difference between a classical and a quantum computer is that the first encodes and processes information on bits, classical entities that can be either 0 or 1, while the second can exploit quantum superposition to encode qubit states of the form $\alpha|0\rangle + \beta|1\rangle$ and entangle multiple of them via unitary gates. But how can such superposition be exploited to have a computational speed-up from a quantum processor?

A common misconception is that such advantage is an immediate consequence of superposition, which would allow a quantum computer to try simultaneously all possible outcomes in a large parallel computation. This erroneous conclusion can probably be traced back to a misinterpretation of the double slit-experiment, which is often explained as if the incident particle passes through both slits simultaneously. What is overlooked here is that the result of the computation is given by the user measuring the final quantum state. When doing so, the user collapses the state to a single outcome, which is randomly generated by Nature according to Born's rule. From a single run of a quantum computer we don't thus obtain multiple results of parallel computations, but a single random outcome. Rather than on the creation of quantum superposition per se, the real quantum computational advantage resides in the clever designing of the interference pattern at the output. In fact, the states associated to different solutions, when encoded in a quantum computer, will have complex amplitudes in the superposition state. The general idea of quantum computing is to design circuits where destructive interference is obtained in the paths which lead to "wrong" answers of the computation, and build up constructing interference for the paths leading to the "right" answer. In this way, the final quantum state will have amplitudes associated to the wrong answers that are close to zero, and the correct answer can then be measured by the user with high-probability. A quantum algorithm can hence be seen as a procedure where gates are used to carefully crafts the probability amplitudes of an input state, so that undesired outputs are suppressed.

In the last few decades an ever growing number of quantum algorithms have been developed to tackle different problems [22, 23]. Some of the most famous ones are Shor's algorithm for prime factoring in polynomial time [24], Grover's algorithm for

polynomial speed-ups in unstructured database search [25], and algorithms for quantum simulation [26–29]. The latter have been described more in details in chapter 6.

A.3.2 *Physical requirements*

One of the principal goals in quantum computing is the development of a universal quantum computer: a computer that can be programmed to implement any conceivable process allowed by quantum mechanics. This would be the most general quantum machine, and represents the Holy Grail for the entire field of quantum information. Via a joint afford of different theoreticians over many years, it has been possible that such a machine could be built in principle [6]. However, in practice, we need to develop a physical hardware to build it. A wide range a platforms are being investigated to achieve that (photons, trapped-ions, superconducting qubits, Majorana fermions, to name a few). The physical requirement that such a platform would need to satisfy to enable universal quantum computation can be summarised in the Di Vincenzo criteria [30]:

- **A scalable physical system with well characterised quantum information carriers.** The single entities encoding quantum information in the platform, e.g. qubits, have to be well defined, and the resources needed to generate and manipulate such states need to grow only at most polynomially with the number of objects used.
- **Initialisation in a simple fiducial state.** The system should be able to produce a fixed initial state to start any computation, which has to be known in advance. A typical initialisation state for qubits is $|00\ldots0\rangle$.
- **Ability to implement a universal gate set.** The architecture needs to be able to support gates for universal quantum computation.
- **Long decoherence times, much longer than the gate operation time.** Interactions with the environment need to be sufficiently low so that the state is preserved during long computations. If this condition is not satisfied, the quantum state would soon decohere into a classical mixed state, almost always leading quantum algorithms to failure.
- **Measurement of the output state.** The platform must allow local measurement on the single quantum information carries to read out the computation result.

An ideal architecture for universal quantum computing needs to satisfy all these requirements simultaneously. Currently, no platform has achieved that yet. The main challenge is to engineer a system that permits interactions strong enough to implement rapid multi-qubit gates, required for universal computation, while at the same time preserving its isolation from the environment.

A.3.3 *Fault tolerance*

In the previous sections we have described how quantum phenomena, such as superposition, manifest themselves at a macroscopic physical scale. In a full-scale quantum computer, however, the hardware would need to support coherent quantum states composed of thousands or millions of qubits, which cannot anymore be globally considered as a microscopic system. A crucial aspect of quantum computing is thus the ability to maintain the quantumness at a macroscopic scale. This requires to fight against the "emergence of classicality" which is naturally observed in macroscopic physical systems. Apart from its many applications, the task of building a large-scale quantum computer is thus also of remarkable fundamental interest. For example, one could conjecture that the reason why the macroscopic world we live in is classical could be due to an underlying fundamental unknown physical law which prohibits quantum superposition to survive in systems above a certain scale, and that quantum computation could therefore be deemed to fail. Although widely considered implausible, such arguments cannot be disregarded a-priori. After all, perpetual motions machines were considered a potentially revolutionary technology by many illustrious scientists, until the discovery of the second law of thermodynamics in the 19th century implied they were impossible [31]. The development of large-scale quantum hardware could thus also bring us new physical insights when starting to explore these uncharted territories.

The field of *quantum error correction* was developed to address these type of concerns from the sceptics of quantum computing. Error correction procedures aim at protecting the computational states from possible errors arising in physical implementations. These can emerge due to unwanted interaction with the environment, or imperfect fabrication or functioning of components. On classical digital computers, where typical rates for computational errors are one every 10^{18} operations, correction of the faulty events can be based on redundancy. As a very simple example, one can think of encoding the state 0 of a bit into n-bit string $000\ldots$ and similarly for 1. If the final logical value is picked as the bit value with highest occurrence in the n-bit string, the possibility of obtaining a wrong value due to small errors gets exponentially suppressed.

Unfortunately, in general such classical redundancy-based schemes cannot be implemented to protect quantum information. The main cause of this is the *no-cloning* theorem, which implies that an arbitrary quantum state $|\psi\rangle$ cannot be deterministically cloned, preventing us to generate copies of an unknown quantum state during the computation unless destroying the original state [32]. More sophisticated techniques are thus required for quantum computing. The first of such quantum error correction protocols was proposed by Shor, where arbitrary errors on a single qubits can be detected and corrected [33]. To do so, the state of each logical qubit was coded in the state of nine entangled qubits. When the computation is performed on qubits encoded in an error

correction code, it is said to be *fault-tolerant*.

As for the classical case, quantum correction is based on adding redundancy in the hardware, with the difference that also entanglement needs to be utilised. Such redundancy implies overheads both in the number of physical qubits and the gates required. These overheads can be in general extremely large. However, the threshold theorem [34] states that, if the noise in individual quantum gates is below a certain threshold, it is possible to efficiently perform an arbitrarily large fault-tolerant quantum computation with only poly-logarithmic additional resources. The implementation of fault-tolerant quantum algorithms, where large quantum states are protected during the computation, can thus be done efficiently. This was a crucial achievement to make quantum computers a realistic prospect.

The required overheads and error rate thresholds remain however quite far from the reach of current quantum platforms. Nonetheless, together with the fast development of more controllable and scalable quantum hardware, many theoretical advances have been achieved in recent years to find fault-tolerant architectures with less stringent demands. The overheads required for modern fault tolerant error correcting schemes, which include stabiliser codes [35, 36] and topological codes [37–39], are orders of magnitude lower then the original proposals.

A.3.4 *Near-term quantum computing machines*

While the development of a fault-tolerant quantum computing architecture remains the main goal for the entire field of quantum information processing, the required overheads prevents the realisation of such machines to be a near-term prospect. On the other hand, in recent years a research topic of increasing interest has been the development of near-term non-universal quantum machines. Such rudimentary platforms can, for example, be composed of tens of non-error-corrected qubits and can perform only algorithms with a limited number of gates. However, while not harnessing the full power of fault-tolerant quantum computation, these rudimentary quantum machines can have already the potential to provide near-term quantum speed-ups in specific applications. These proposals include *boson samplers* (see chapter 4) [40], sampling from instantaneous quantum polynomial time (IQP) circuits [41], and sampling the output state of a system evolving under nearest-neighbour, translationally-invariant Hamiltonians [42]. These proposals are now being pursued by many research group and companies around the globe, and a near-term demonstration of a quantum speed-up using these machines is expected to happen in the next few years.

References

1. Barber, D. *Bayesian reasoning and machine learning* (Cambridge University Press, 2012).
2. Zeilinger, A., Gähler, R., Shull, C. G., Treimer, W. & Mampe, W. Single- and double-slit diffraction of neutrons. *Rev. Mod. Phys.* **60,** 1067–1073. https://link.aps.org/doi/10.1103/RevModPhys.60.1067 (1988).
3. Arndt, M., Nairz, O., Vos-Andreae, J., Keller, C., Van der Zouw, G. & Zeilinger, A. Wave–particle duality of C_{60} molecules. *Nature* **401,** 680 (1999).
4. Eibenberger, S., Gerlich, S., Arndt, M., Mayor, M. & Tüxen, J. Matter–wave interference of particles selected from a molecular library with masses exceeding 10000 amu. *Physical Chemistry Chemical Physics* **15,** 14696–14700 (2013).
5. Arndt, M. & Hornberger, K. Testing the limits of quantum mechanical superpositions. *Nature Physics* **10,** 271 (2014).
6. Nielsen, M. A. & Chuang, I. *Quantum computation and quantum information* (Cambridge University Press, 2002).
7. Einstein, A., Podolsky, B. & Rosen, N. Can Quantum-Mechanical Description of Physical Reality Be Considered Complete? *Phys. Rev.* **47,** 777. http://link.aps.org/doi/10.1103/PhysRev.47.777 (1935).
8. Schrödinger, E. Discussion of Probability Relations between Separated Systems. *Proc. Cambridge Philos. Soc.* **51,** 553 (1935).
9. Horodecki, R., Horodecki, P., Horodecki, M. & Horodecki, K. Quantum entanglement. *Rev. Mod. Phys.* **81,** 865–942. http://link.aps.org/doi/10.1103/RevModPhys.81.865 (2009).
10. Amico, L., Fazio, R., Osterloh, A. & Vedral, V. Entanglement in many-body systems. *Rev. Mod. Phys.* **80,** 517–576. http://link.aps.org/doi/10.1103/RevModPhys.80.517 (2008).
11. Werner, R. F. Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model. *Phys. Rev. A* **40,** 4277–4281. http://link.aps.org/doi/10.1103/PhysRevA.40.4277 (1989).
12. Plenio, M. B. & Virmani, S. An introduction to entanglement measures. *Quantum Information & Computation* **7,** 1–51 (2007).
13. Bell, J. S. On the Einstein-Podolsky-Rosen Paradox. *Physics* **1,** 195 (1964).
14. Clauser, J. F., Horne, M. A., Shimony, A. & Holt, R. A. Proposed Experiment to Test Local Hidden-Variable Theories. *Phys. Rev. Lett.* **23,** 880–884. https://link.aps.org/doi/10.1103/PhysRevLett.23.880 (1969).
15. Aspect, A., Grangier, P. & Roger, G. Experimental Tests of Realistic Local Theories via Bell's Theorem. *Phys. Rev. Lett.* **47,** 460. http://link.aps.org/doi/10.1103/PhysRevLett.47.460 (1981).

16. Hensen, B., Bernien, H., Dréau, A. E., Reiserer, A., Kalb, N., Blok, M. S., Ruitenberg, J., Vermeulen, R. F., Schouten, R. N., Abellán, C., *et al.* Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres. *Nature* **526,** 682 (2015).

17. Shalm, L. K. *et al.* Strong Loophole-Free Test of Local Realism. *Phys. Rev. Lett.* **115,** 250402. https://link.aps.org/doi/10.1103/PhysRevLett.115.250402 (2015).

18. Giustina, M. *et al.* Significant-Loophole-Free Test of Bell's Theorem with Entangled Photons. *Phys. Rev. Lett.* **115,** 250401. https://link.aps.org/doi/10.1103/PhysRevLett.115.250401 (2015).

19. Cirel'son, B. S. Quantum generalizations of Bell's inequality. *Letters in Mathematical Physics* **4,** 93–100 (1980).

20. Brunner, N., Cavalcanti, D., Pironio, S., Scarani, V. & Wehner, S. Bell nonlocality. *Rev. Mod. Phys.* **86,** 419–478. https://link.aps.org/doi/10.1103/RevModPhys.86.419 (2014).

21. Ladd, T. D., Jelezko, F., Laflamme, R., Nakamura, Y., Monroe, C. & O'Brien, J. L. Quantum computers. *Nature* **464,** 45 (2010).

22. *Quantum Algorithm Zoo.* https://math.nist.gov/quantum/zoo/.

23. Montanaro, A. Quantum algorithms: an overview. *npj Quantum Information* **2,** 15023 (2016).

24. Shor, P. Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer. *SIAM J. Comput.* **26,** 1484 (1997).

25. Grover, L. K. A fast quantum mechanical algorithm for database search. *Proceedings of the twenty-eighth annual ACM symposium on Theory of computing,* 212–219 (1996).

26. Feynman, R. Simulating Physics with Computers. *Int. J. Theor. Phys.* **21,** 467 (1982).

27. Lloyd, S. Universal quantum simulators. *Science,* 1073–1078 (1996).

28. Huh, J., Guerreschi, G. G., Peropadre, B., McClean, J. R. & Aspuru-Guzik, A. Boson sampling for molecular vibronic spectra. *Nature Photonics* **9,** 615 (2015).

29. Peruzzo, A., McClean, J., Shadbolt, P., Yung, M.-H., Zhou, X.-Q., Love, P. J., Aspuru-Guzik, A. & O'Brien, J. L. A variational eigenvalue solver on a photonic quantum processor. *Nature communications* **5,** 4213 (2014).

30. DiVincenzo, D. P. The physical implementation of quantum computation. *Fortschritte der Physik* **48,** 771–783 (2000).

31. *Perpetual Motion of The 21st Century?* https://rjlipton.wordpress.com/2012/01/30/perpetual-motion-of-the-21st-century/. Blog post.

32. Wootters, W. K. & Zurek, W. H. A single quantum cannot be cloned. *Nature* **299,** 802–803 (1982).

33. Shor, P. W. Scheme for reducing decoherence in quantum computer memory. *Physical review A* **52,** R2493 (1995).

34. Aharonov, D. & Ben-Or, M. Fault-tolerant quantum computation with constant error rate. *arXiv:quant-ph/9906129* (1999).

35. Gottesman, D. Class of quantum error-correcting codes saturating the quantum Hamming bound. *Phys. Rev. A* **54,** 1862–1868. https://link.aps.org/doi/10.1103/PhysRevA.54.1862 (1996).

36. Calderbank, A. R., Rains, E. M., Shor, P. W. & Sloane, N. J. A. Quantum Error Correction and Orthogonal Geometry. *Phys. Rev. Lett.* **78,** 405–408. https://link.aps.org/doi/10.1103/PhysRevLett.78.405 (1997).

37. Kitaev, A. Y. in *Quantum Communication, Computing, and Measurement* 181–188 (Springer, 1997).

38. Dennis, E., Kitaev, A., Landahl, A. & Preskill, J. Topological quantum memory. *Journal of Mathematical Physics* **43,** 4452–4505 (2002).

39. Raussendorf, R., Harrington, J. & Goyal, K. Topological fault-tolerance in cluster state quantum computation. *New Journal of Physics* **9,** 199 (2007).

40. Aaronson, S. & Arkhipov, A. The computational complexity of linear optics. *Quant. Inf. Comp.* **13,** 1383 (2014).

41. Bremner, M. J., Montanaro, A. & Shepherd, D. J. Achieving quantum supremacy with sparse and noisy commuting quantum computations. *Quantum* **1,** 8 (2017).

42. Bermejo-Vega, J., Hangleiter, D., Schwarz, M., Raussendorf, R. & Eisert, J. Architectures for Quantum Simulation Showing a Quantum Speedup. *Phys. Rev. X* **8,** 021010. https://link.aps.org/doi/10.1103/PhysRevX.8.021010 (2018).

B REVERSE HONG-OU-MANDEL EFFECT IN THE GENERAL CASE

In this appendix we provide a detailed derivation of eq.(2.2.15), which gives the general form of the rev-HOM interference in presence of different photon-pair sources. The merit of it is to formally show that the visibility of the rev-HOM effect is directly linked to overlap of the joint spectra, and that the rev-HOM can thus be used to characterise the indistinguishability of photon sources. Indeed, while the fact that the rev-HOM represents a characterisation tool is often claimed in the literature, a formal derivation of why is that and what exactly it characterises does not appear, to the author's knowledge, to have been explicitly reported.

B.1 Reverse Hong-Ou-Mandel with arbitrary spectra

In the rev-HOM effect, we inject a superposition state $|\psi\rangle = (|\psi\rangle_1 + e^{i2\phi}\,|\psi\rangle_2)/\sqrt{2}$ into a beam-splitter (see Fig. 2.3), where $|\psi\rangle_1$ and $|\psi\rangle_2$ are the biphoton states associated to the first and second input modes of the beam-splitter, respectively, which we can write as:

$$|\psi\rangle_1 = \int d\omega_i d\omega_s \; \psi_1(\omega_i,\omega_s)\hat{a}_1^\dagger(\omega_i)\hat{a}_1^\dagger(\omega_s)\,|0\rangle\,, \tag{B.1.1}$$

$$|\psi\rangle_2 = \int d\omega_i d\omega_s \; \psi_2(\omega_i,\omega_s)\hat{a}_2^\dagger(\omega_i)\hat{a}_2^\dagger(\omega_s)\,|0\rangle\,. \tag{B.1.2}$$

Here, $\psi_1(\omega_i,\omega_s)$ and $\psi_2(\omega_i,\omega_s)$ are the two complex joint spectra for the biphoton state on each mode. The action the beam-splitter, which, for simplicity, is supposed to be

frequency-independent, can be written as the transformation of the bosonic operators

$$\hat{a}_1^\dagger(\omega) \mapsto \left[\hat{a}_1^\dagger(\omega) + i\hat{a}_2^\dagger(\omega)\right]/\sqrt{2}, \tag{B.1.3}$$

$$\hat{a}_2^\dagger(\omega) \mapsto \left[i\hat{a}_1^\dagger(\omega) + \hat{a}_2^\dagger(\omega)\right]/\sqrt{2}. \tag{B.1.4}$$

The output state is thus given by

$$|\psi\rangle \mapsto \quad \frac{1}{2}\int d\omega_i d\omega_s\ \psi_1(\omega_i,\omega_s)\left[\left(\hat{a}_1^\dagger(\omega_i) + i\hat{a}_2^\dagger(\omega_i)\right)\left(\hat{a}_1^\dagger(\omega_s) + i\hat{a}_2^\dagger(\omega_s)\right)\right]|0\rangle \tag{B.1.5}$$

$$+e^{i2\phi}\psi_2(\omega_i,\omega_s)\left[\left(i\hat{a}_1^\dagger(\omega_i) + \hat{a}_2^\dagger(\omega_i)\right)\left(i\hat{a}_1^\dagger(\omega_s) + \hat{a}_2^\dagger(\omega_s)\right)\right]|0\rangle \tag{B.1.6}$$

$$= \quad |\text{bunch}\rangle + |\text{split}\rangle \tag{B.1.7}$$

where

$$|\text{bunch}\rangle = \frac{1}{2}\int d\omega_i d\omega_s \left[\hat{a}_2^\dagger(\omega_i)\hat{a}_2^\dagger(\omega_s) - \hat{a}_1^\dagger(\omega_i)\hat{a}_1^\dagger(\omega_s)\right]\left(e^{i2\phi}\psi_2(\omega_i,\omega_s) - \psi_1(\omega_i,\omega_s)\right)|0\rangle \tag{B.1.8}$$

$$|\text{split}\rangle = \frac{i}{2}\int d\omega_i d\omega_s \left[\hat{a}_1^\dagger(\omega_i)\hat{a}_2^\dagger(\omega_s) + \hat{a}_2^\dagger(\omega_i)\hat{a}_1^\dagger(\omega_s)\right]\left(e^{i2\phi}\psi_2(\omega_i,\omega_s) + \psi_1(\omega_i,\omega_s)\right)|0\rangle\,. \tag{B.1.9}$$

Here, $|\text{bunch}\rangle$ represents the cases where the output photons bunch in the same mode, while the the term $|\text{split}\rangle$ represents the cases where photons emerge on different modes. If our measurements are frequency insensitive (i.e. the detectors do not resolve difference wavelengths), than the single-photon detection projective operators on the two different output modes are given by:

$$\hat{P}_1 = \int d\omega_1\ \hat{a}_1^\dagger(\omega_1)\,|0\rangle\,\langle 0|\,\hat{a}_1(\omega_1), \tag{B.1.10}$$

$$\hat{P}_2 = \int d\omega_2\ \hat{a}_2^\dagger(\omega_2)\,|0\rangle\,\langle 0|\,\hat{a}_2(\omega_2), \tag{B.1.11}$$

and the projective operator for a coincidence measurements, i.e. both detectors on modes 1 and 2 click, is given by

$$\hat{P}_1 \otimes \hat{P}_2 = \int d\omega_1 d\omega_2\ \hat{a}_1^\dagger(\omega_1)\hat{a}_2^\dagger(\omega_2)\,|0\rangle\,\langle 0|\,\hat{a}_1(\omega_1)\hat{a}_2(\omega_2). \tag{B.1.12}$$

The coincidence probability is thus given by the expectation value $p_{\text{coinc}} = \langle\psi|\,\hat{P}_1 \otimes \hat{P}_2\,|\psi\rangle$. Now, the term $|\text{bunch}\rangle$, which only contains terms $\hat{a}_1^\dagger\hat{a}_1^\dagger$ or $\hat{a}_2^\dagger\hat{a}_2^\dagger$, necessarily gives zero

contribution to this expectation value. We thus have

$$p_{\text{coinc}} = \langle \text{split}| \hat{P}_1 \otimes \hat{P}_2 |\text{split}\rangle = \frac{1}{4}\int d\omega_i d\omega_s d\omega_1 d\omega_2 d\omega_i' d\omega_s' \tag{B.1.13}$$

$$\left[e^{-i2\phi}\psi_2^*(\omega_i,\omega_s) + \psi_1^*(\omega_i,\omega_s)\right]\left[e^{i2\phi}\psi_2(\omega_i',\omega_s') + \psi_1(\omega_i',\omega_s')\right] \tag{B.1.14}$$

$$\times \langle 0|\left[\hat{a}_1(\omega_i)\hat{a}_2(\omega_s)\hat{a}_1^\dagger(\omega_1)\hat{a}_2^\dagger(\omega_2) + \hat{a}_2(\omega_i)\hat{a}_1(\omega_s)\hat{a}_1^\dagger(\omega_1)\hat{a}_2^\dagger(\omega_2)\right]|0\rangle \tag{B.1.15}$$

$$\times \langle 0|\left[\hat{a}_1(\omega_1)\hat{a}_2(\omega_2)\hat{a}_1^\dagger(\omega_i')\hat{a}_2^\dagger(\omega_s') + \hat{a}_1(\omega_1)\hat{a}_2(\omega_2)\hat{a}_2^\dagger(\omega_i')\hat{a}_1^\dagger(\omega_s')\right]|0\rangle \tag{B.1.16}$$

$$= \frac{1}{4}\int d\omega_1 d\omega_2 \tag{B.1.17}$$

$$\left[e^{-i2\phi}\psi_2^*(\omega_1,\omega_2) + \psi_1^*(\omega_1,\omega_2)\right]\left[e^{i2\phi}\psi_2(\omega_1,\omega_2) + \psi_1(\omega_1,\omega_2)\right] \tag{B.1.18}$$

$$+ \left[e^{-i2\phi}\psi_2^*(\omega_1,\omega_2) + \psi_1^*(\omega_1,\omega_2)\right]\left[e^{i2\phi}\psi_2(\omega_2,\omega_1) + \psi_1(\omega_2,\omega_1)\right] \tag{B.1.19}$$

$$+ \left[e^{-i2\phi}\psi_2^*(\omega_2,\omega_1) + \psi_1^*(\omega_2,\omega_1)\right]\left[e^{i2\phi}\psi_2(\omega_1,\omega_2) + \psi_1(\omega_1,\omega_2)\right] \tag{B.1.20}$$

$$+ \left[e^{-i2\phi}\psi_2^*(\omega_2,\omega_1) + \psi_1^*(\omega_2,\omega_1)\right]\left[e^{i2\phi}\psi_2(\omega_2,\omega_1) + \psi_1(\omega_2,\omega_1)\right] \tag{B.1.21}$$

$$= \frac{1}{2} + \frac{1}{2}\,\text{Re}\left[e^{i2\phi}\int d\omega_1 d\omega_2\ \psi_1^*(\omega_1,\omega_2)\psi_2(\omega_1,\omega_2)\right], \tag{B.1.22}$$

where we have used $\langle 0|\,\hat{a}_1(\omega_i)\hat{a}_2(\omega_s)\hat{a}_1^\dagger(\omega_1)\hat{a}_2^\dagger(\omega_2)\,|0\rangle = \delta(\omega_i,\omega_1)\delta(\omega_s,\omega_2)$, and the symmetry $\psi(\omega_1,\omega_2) = \psi(\omega_2,\omega_1)$ and normalisation $\int d\omega_1 d\omega_2 |\psi(\omega_1,\omega_2)|^2 = 1$ of the joint spectra. Defining $\theta_0 = \frac{1}{2}\arg\int d\omega_1 d\omega_2 \psi_2^*(\omega_1,\omega_2)\psi_1(\omega_1,\omega_2)$, we finally arrive at eq.(**??**)

$$p_{\text{coinc}} = \frac{1}{2} + \frac{1}{2}\cos(2\phi+\theta_0)\left|\int d\omega_1 d\omega_2\ \psi_1^*(\omega_1,\omega_2)\psi_2(\omega_1,\omega_2)\right| \tag{B.1.23}$$

which represents the rev-HOM fringe. As discussed in section 2.2.2, the visibility

$$\mathcal{V}_{\text{rev-HOM}} = \left|\int d\omega_1 d\omega_2\ \psi_2^*(\omega_1,\omega_2)\psi_1(\omega_1,\omega_2)\right| \tag{B.1.24}$$

is directly linked to the overlap between the two joint spectra, but does not depend on their separability, in contrast to the standard HOM effect.

C METHODS FOR CHARACTERISING HIGH-DIMENSIONAL ENTANGLEMENT

In this appendix we provide the theoretical background and some of the details for the applications investigated in the experiment in chapter 3. Contributions to the material described here were provided by Laura Mančinska for the dimension witness section, Alexia Salavrakos, Jordi Tura, Remigiusz Augusiak and Antonio Acín for the SATWAP inequalities section, and Paul Skrzypczyk for the section on steering.

C.1 Dimension witness

Dimension witnesses are protocols for the certification of the local dimensionality of a system. While different approaches are possible, here we follow the one developed in Ref. [1]. In particular, the bounds on the local dimensions are obtained using the following theorem:

THEOREM 1 ([1]). Suppose a correlation $\mathbf{p} = p(ab|xy)$ is produced by two separated parties Alice and Bob who do not have access to shared randomness but possess a shared quantum state $|\psi\rangle$ on $\mathbb{C}^{d_1} \otimes C^{d_2}$. Then

$$\lceil \mathcal{D}(\mathbf{p}) \rceil = \left\lceil \frac{1}{f(\mathbf{p})} \right\rceil \leq \min\{d_1, d_2\}$$

where

$$f(\mathbf{p}) := \min_{y,y'} \sum_{b,b'} \min_{x} \left(\sum_{a} \sqrt{p(ab|xy)p(ab'|xy')} \right)^2 . \tag{C.1.1}$$

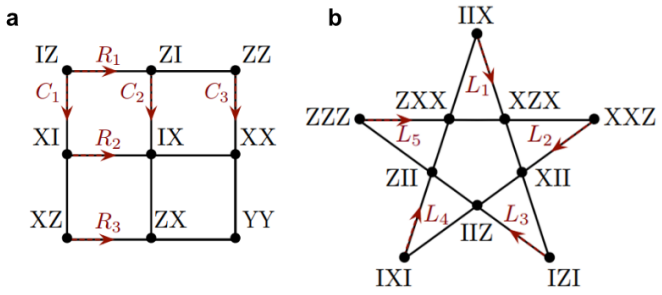


Figure C.1: Correlation measurements associated to optimal strategies for Magic Square and Pentagram games. X, Y and Z are Pauli operators and I is the identity. Red lines C_i, R_i and L_i are associated to different measurement settings. Single Magic Square game, Magic Pentagram game, and two copies of the Magic Square game are used in the experiment in chapter 3 to certify the dimension for states with local dimension $d = 4$, $d = 6, 8$, and $d = 10, 12, 14, 15$, respectively.

To give an upper bound to $f(\mathbf{p})$, and therefore a lower bound to the local dimension, instead of minimising over all possible values y, y', and x in Eq. C.1.1 we can do measurements for some fixed measurement settings, y, y' for Bob and x for Alice, and calculate the function

$$f_{y,y',x}(\mathbf{p}) := \sum_{b,b'} \left(\sum_{a} \sqrt{p(ab|xy)p(ab'|xy')} \right)^2 .$$

Indeed, for any choice of settings, y, y' and x the value $f_{y,y',x}(\mathbf{p})$ gives an upper bound the optimised value $f(\mathbf{p})$. Therefore, the inverse, $1/f_{y,y',x}(\mathbf{p})$, lower bounds the local dimensions $\min\{d_1, d_2\}$. In the cases considered in the experiment of chapter 3, we can choose y, y', x so that $f_{y,y',x}(\mathbf{p}) = f(\mathbf{p})$, as shown below. Therefore, instead of collecting the full correlation data for all possible measurement settings, it is sufficient to fix the measurement setting for Alice and obtain the data for two possible measurement settings for Bob. Note that, independently of the correlation $\mathbf{p}$, it will always be possible to only collect measurement statistics for two different measurement settings for Bob by picking two values of y and y' which attain the minimum in Equation C.1.1. On the other hand, we might need to use many different measurement settings for Alice in order to obtain the best bound. This is due to the fact that Alice's setting x attaining the minimum in Equation C.1.1 could depend on b and b'.

Magic Square correlations. In the Magic Square game [2–4] (see Fig. C.1a) each of Alice and Bob have three different measurement settings: $\{1, 2, 3\}$. Alice's outputs are $A = \{a \in \{+1, -1\}^3 : a_1 a_2 a_3 = 1\}$ and Bob's outputs are $B = \{b \in \{+1, -1\}^3 : b_1 b_2 b_3 =$

$-1\}$. The players win the game, if their outputs, a and b, satisfy $a_y = b_x$, where x and y are the measurement settings. Although the Magic Square game cannot be won using classical correlations, the following quantum correlation $\mathbf{p}$ allows us to win independently from the selected measurement settings:

$$\mathbf{p} = p(ab|xy) = \begin{cases} \frac{1}{8} & \text{if } a_y = b_x \\ 0 & \text{otherwise} \end{cases} \tag{C.1.2}$$

This correlation can be realised by measurements on a bipartite maximally entangled state of local dimension 4. In particular, the ith measurement setting for Bob (Alice) corresponds to measuring in the orthonormal basis diagonalising all the operators in the ith column (row) in the square grid shown in Fig. C.1a, where the 4-dimensional local Hilbert space is represented as a 2-qubit space. The outcome Bob (Alice) obtains corresponds to a simultaneous eigenvector of the the three operators in the respective column (row) with eigenvalues $\lambda_1, \lambda_2, \lambda_3 \in \{+1, -1\}$. Since the along any row (column) the operators multiply to I $(-I)$. The eigenvalues obtained by Alice will multiply to $+1$, while those obtained by Alice will multiply to -1. Therefore, the eigenvalue strings $(\lambda_1, \lambda_2, \lambda_3)$ obtained by each of the players correspond to valid answers.

Using $y = x = 1$ and $y' = 2$ we find that

$$f_{y,y',x}(\mathbf{p}) = \sum_{b,b' \in B} \left(\sum_{a \in A} \sqrt{p(ab|11)p(ab'|12)} \right)^2 = \sum_{b,b' \in B} \left(\sqrt{\frac{1}{8} \cdot \frac{1}{8}} \right)^2 = \frac{|B|^2}{64} = \frac{16}{64} = \frac{1}{4} \tag{C.1.3}$$

and therefore we can certify that the local dimension of the measured system is at least four. Note that these conclusions are obtained solely from observing correlation $\mathbf{p}$ and without making any assumption about the performed measurements or the shared state, though we do need the assumption that shared randomness is not free. The above analysis applies to the ideal case with no experimental error. In an actual experiment we can gather outcome statistics to obtain probabilities, $\hat{p}(ab|xy)$. If the experiment has low noise, the correlation $\hat{\mathbf{p}}$ will be very close to $\mathbf{p}$ and we will get $\lceil 1/f_{y=1,y'=2,x=1}(\hat{\mathbf{p}}) \rceil = 4$ thus certifying that the dimension of the measured system is 4. In particular, in order to evaluate $f_{y=1,y'=2,x=1}$ we can decide to gather measurement statistics for measurement settings $y = 1$ and $y' = 2$ for Bob and $x = 1$ for Alice. According to Fig. C.1a, this amounts to measuring both of Alice's qubits in the Z-basis $(x = 1)$ and either measuring Bob's first qubit in the X-basis and the second one in the Z-basis $(y = 1)$ or the other way around $(y' = 2)$, where X, Y, and Z are the Pauli matrices. In the experiment in chapter 3, we refer to this case as "scenario I" and denote the experimentally observed distribution by $\mathbf{p}_4^{\mathrm{I}}$.

We can also think of playing two copies of the Magic Square game in parallel, where the players win if the winning conditions for both games are satisfied simultaneously. The correlation $\mathbf{p_2}$ that lets us win this game can be realized by measurements on a bipartite maximally entangled state of local dimension 16, where each player holds the equivalent of a 4-qubit system. Via a calculation similar to the one in C.1.3, we find that $f_{y=(1,1),y'=(2,2),x=(1,1)}(\mathbf{p_2}) = 1/16$, and hence correlation $\mathcal{D}(\mathbf{p}) \geq 16$ and this correlation certifies dimension 16.

Magic Pentagram correlations. In the Magic Pentagram game [2] (see Fig. C.1b) each of Alice and Bob have five different measurement settings: $\{1, 2, 3, 4, 5\}$ and we think of each of these settings as identifying one the five lines in the pentagram in Fig. C.1b in the main text. Alice's and Bob's outputs are $C^+ = \{c \in \{+1, -1\}^4 : c_1c_2c_3c_4 = 1\}$ for settings 2 to 4 and their outputs are $C^- = \{c \in \{+1, -1\}^4 : c_1c_2c_3c_4 = -1\}$ for setting 5. We think of the ith output position as assigning value to the ith point on the line corresponding to the measurement setting. For measurement settings x and y, the players win the game, if their outputs, a and b, assign the same value to all the points belonging to both lines L_x and L_y.

For instance if we ask Bob the fifth line ($y = 5$) and Alice the third one ($x = 3$) then these two lines intersect in precisely one point (the 1st point on L_5 and the 5th point on L_3). So to win, the first position in Bob's answer string should equal the last position in Alice's answer string; for example, $b = (+1, -1, +1, +1)$ and $a = (+1, -1, -1, +1)$ would be a winning answer pair.

Just like in the case of Magic Square, there exists a quantum correlation $\mathbf{p}$ that can be used to win the Magic Pentagram game for all measurement settings but no classical correlation would achieve this. Specifically,

$$\mathbf{p} = p(ab|xy) = \begin{cases} \frac{1}{32} & \text{if } a \text{ and } b \text{ satisfy conditions (1) and (2) above} \\ 0 & \text{otherwise} \end{cases} \tag{C.1.4}$$

and it can be realized by measurements on a maximally entangled state of local dimension 8, locally equivalent to a 3-qubit space. In particular, the ith measurement setting for Bob (Alice) corresponds to measuring in the orthonormal basis diagonalizing all the commuting observables on the ith line of Fig. C.1b. The outcome corresponds to a simultaneous eigenvector of all four operators and the players respond with a string of eigenvalues $(\lambda_1, \ldots, \lambda_4)$ that correspond to their measurement outcome. This string will be a valid answer since the observables on the first line multiply to $-I$ and to I along all the other lines.

Fixing $y = x = 3$ and $y' = 5$, we find that

$$f_{y,y',x}(\mathbf{p}) = \sum_{b\in C^+, b'\in C^-} \left(\sum_{a\in C^+} \sqrt{p(ab|33)p(ab'|35)} \right)^2$$
$$= \sum_{b\in C^+, b'\in C^-} p(bb|33)p(bb'|35) = \sum_{b\in C^+} \frac{1}{32} \sum_{b'\in C^-} p(bb'|35) = \frac{|C^+| \cdot 8}{32 \cdot 32} = \frac{1}{8} \tag{C.1.5}$$

as $p(ab|33) = \delta_{ab}/32$ and $p(bb'|35) = \delta_{b_5 b'_1}/32$. It now follows that we can certify that the local dimension of the measured system is at least 8 by observing correlation $\mathbf{p}$.

Scenario I. The measurements we perform in this scenario are based on the optimal strategies for Magic Square and Magic Pentagram games that we explained above. In the experiment in chapter 3, for dimensions $d = 4, 8$ we prepared the d-dimensional maximally entangled state $\left|\psi_d^+\right\rangle = \frac{1}{\sqrt{d}} \sum_{k=0}^{d-1} |k,k\rangle$ and then perform the measurements corresponding to optimal strategies for Magic Square game ($d = 4$) or Magic Pentagram game ($d = 8$). Instead of gathering statistics for all possible measurement settings in these games we chose only one setting for Alice (setting 1 for Magic Square and setting 3 for Magic Pentagram) and two different settings for Bob (settings 1, 2 for Magic Square and settings 3, 5 for Magic Pentagram). We argued above that in the absence of experimental error, these correlations would enable us to certify dimensions 4 and 8 respectively. Experimental results are reported in Fig. 3.7.

In order to certify intermediate dimensions $4 < d < 8$ we prepared the state

$$\left|\hat{\psi}_d^+\right\rangle = \frac{1}{\sqrt{d}} \sum_{k=0}^{d-1} |k,k\rangle \tag{C.1.6}$$

where $|k\rangle \in \mathbb{C}^8$. Note that although essentially $\left|\hat{\psi}_d^+\right\rangle$ is the d-dimensional maximally entangled state, it is embedded in dimension $\mathbb{C}^8 \otimes \mathbb{C}^8$ so we can measure it using the 8-dimensional basis measurements from the optimal strategy of Magic Pentagram game. As before, we only use measurement setting 3 for Alice and settings 3 and 5 for Bob.

In order to certify even dimensions $8 < d < 16$ we used a similar approach but performed measurements from the optimal strategy for two copies of Magic Square instead of Magic Pentagram game. Specifically, we prepared the state $\left|\hat{\psi}_d^+\right\rangle = \frac{1}{\sqrt{d}} \sum_{k=0}^{d-1} |k,k\rangle$ where $|k\rangle \in \mathbb{C}^{16}$, and measured it using setting $(1,1)$ for Alice and settings $(1,1)$ and $(2,2)$ for Bob. For Alice this amounts to measuring her state in the Z-basis while Bob either measures his first and third qubit in the X-basis and the remaining two in the Z-basis (setting (1,1)) or the other way around (setting (2,2)).

Scenario II. In this scenario, to obtain the correlation $\mathbf{p}_d^{\mathrm{II}}$, we prepare the d-dimensional maximally entangled state $\left|\psi_d^+\right\rangle$ for $4 \leq d \leq 15$ and measure each of the two local systems in the $\hat{Z}$-basis. In the ideal case, we would obtain a correlation $p_d(ab|11) = \delta_{ab}/d$, where $a, b \in \{0, \ldots, d-1\}$. Although a similar conclusion could also be obtained in a more straightforward manner, we can compute the bound $\mathcal{D}(\mathbf{p}_d)$ to see which dimension would be certified by experimentally observing this correlation. Since there is only a single measurement setting for both Alice and Bob (i.e., a single choice for the values y, y' and x), we see that

$$\begin{aligned} f(\mathbf{p}_d) &= \sum_{b,b'} \left(\sum_a \sqrt{p(ab|11)p(ab'|11)} \right)^2 \\ &= \sum_{b,b'} \left(\sum_a \frac{\delta_{ab}}{\sqrt{d}} \sqrt{p(ab'|11)} \right)^2 = \frac{1}{d} \sum_{b,b'} p(bb'|11) = \frac{1}{d} \end{aligned} \tag{C.1.7}$$

as the only non-zero term in the sum over a is the one where $a = b$. Therefore, it follows that $\mathcal{D}(\mathbf{p}_d) = 1/f(\mathbf{p}_d) = d$ and the correlation $\mathbf{p}_d$ can be used to certify dimension d. Of course, experimentally we do not observe exactly $\mathbf{p_d}$. Experimental results are reported in Fig. 3.7.

C.2 Bell inequalities for multidimensional bipartite maximally entangled states.

Here we describe in more details the class of Bell inequalities used in chapter 3 to test non-locality of the maximally entangled states of two qudits, which were recently proposed [5]. Here we follow the derivation obtained in Ref. [5] Let us consider a Bell scenario in which two parties A and B share some quantum state $\hat{\rho}$. We then assume that each party performs one of two measurements, $A_x = \{\hat{M}_{a|x}\}_a$ for A and $B_y = \{\hat{M}_{b|y}\}_b$ for B with $x, y = 1, 2$. Each measurement yields one of d possible outcomes, which are labelled by $0, \ldots, d-1$.

Such local measurements, after many repetitions, lead to correlations that are described by a collection of joint probabilities

$$\{p(ab|xy)\}_{a,b \in \{0,\ldots,d-1\}} \tag{C.2.1}$$

with $x, y = 1, 2$, where each $p(ab|xy) \equiv p(A_x = a, B_y = b)$ is the probability that A and B obtain outcomes a and b upon performing the measurements A_x and B_y, respectively, and can be expressed by the Born rule as $p(ab|xy) = \mathrm{Tr}[\hat{\rho}(\hat{M}_{a|x} \otimes \hat{M}_{b|y})]$ (see Fig. 3.6).

The class of Bell inequalities used to test non-locality of the maximally entangled states is [5]:

$$I_d := \sum_{k=0}^{\lfloor d/2 \rfloor - 1} (\alpha_k \mathbb{P}_k - \beta_k \mathbb{Q}_k) \leq C_d, \tag{C.2.2}$$

where the expressions $\mathbb{P}_k$ and $\mathbb{Q}_k$ are defined as

$$\mathbb{P}_k = P(A_1 = B_1 + k) + P(B_1 = A_2 + k) + P(A_2 = B_2 + k) + P(B_2 = A_1 + k + 1) \tag{C.2.3}$$

and

$$\mathbb{Q}_k = P(A_1 = B_1 - k - 1) + P(B_1 = A_2 - k - 1) + P(A_2 = B_2 - k - 1) + P(B_2 = A_1 - k). \tag{C.2.4}$$

The coefficients α_k and β_k are given by

$$\alpha_k = \frac{1}{2d}\left[g(k) + (-1)^d \tan\left(\frac{\pi}{4d}\right)\right], \qquad \beta_k = \frac{1}{2d}\left[g\left(k + 1/2\right) - (-1)^d \tan\left(\frac{\pi}{4d}\right)\right], \tag{C.2.5}$$

with $g(k) = \cot[\pi(k + 1/4)/d]$. Here, $P(A_x = B_y + k)$ is the probability that upon measuring A_x and B_y, the results obtained by Alice and Bob differ by k mod d, that is

$$P(A_x = B_y + k) = \sum_{a=0}^{d-1} p((a + k \bmod d)a|xy). \tag{C.2.6}$$

It is now convenient to rewrite the inequality of (C.2.2) in the correlator form. As we work with more than two outcomes, it useful to introduce the notion of generalised complex correlators, which are defined as the two-dimensional Fourier transform of the joint distributions (C.2.1),

$$\langle A_x^k B_y^l \rangle = \sum_{a,b=0}^{d-1} \omega^{ak+bl} p(ab|xy), \tag{C.2.7}$$

for $x, y = 1, 2$ and $k, l = 0, \ldots, d-1$. Here $\omega = \exp(2\pi i/d)$ is a root of the unity. Notice that when $d = 2$, the definition of the correlators $\langle A_x B_y \rangle = P(A_x = B_y) - P(A_x \neq B_y)$ naturally also corresponds to taking the discrete Fourier transform. Expressing the probabilities $p(ab|xy)$ in terms of the generalised correlators and then plugging into (C.2.2), we obtain

$$\begin{aligned} \widetilde{I}_d &:= \sum_{l=1}^{d-1} \left[a_l \langle A_1^l B_1^{d-l} \rangle + a_l^* \omega^l \langle A_1^l B_2^{d-l} \rangle + a_l \langle A_2^l B_2^{d-l} \rangle + a_l^* \langle A_2^l B_1^{d-l} \rangle\right] \\ &= 2 \sum_{l=1}^{\lfloor d/2 \rfloor - 1} \mathrm{Re}\left[\langle A_1^l \bar{B}_1^l \rangle + \langle A_2^l \bar{B}_2^l \rangle\right] \leq \mathcal{C}_d, \end{aligned} \tag{C.2.8}$$

where $\mathcal{C}_d$ is its classical bound given by

$$\mathcal{C}_d = \frac{1}{2}\left[3\cot\left(\frac{\pi}{4d}\right) - \cot\left(\frac{3\pi}{4d}\right)\right] - 2, \tag{C.2.9}$$

and $\bar{B}_i^l$ are new variables defined as $\bar{B}_1^l = a_l B_1^{d-l} + a_l^* \omega^l B_2^{d-l}$ and $\bar{B}_2^l = a_l B_2^{d-l} + a_l^* B_1^{d-l}$ with $a_l = \omega^{(2l-d)/8}/\sqrt{2}$. Notice that in the quantum case A_x and B_y are unitary operators of eigenvalues $1, \omega, \ldots, \omega^{d-1}$, whereas A_x^k and B_y^l are simply their matrix powers. Then, the complex correlators can be easily expressed with the aid of the Born rule as $\langle A_x^k B_y^l \rangle = \mathrm{Tr}\left[\hat{\rho}(A_x^k \otimes B_y^l)\right]$. We thus can think of A_x and B_y as quantum observables whose results are for our convenience labelled by the roots of unity $1, \omega, \ldots, \omega^{d-1}$.

The maximal quantum violation of the inequality (C.2.8) is $\mathcal{Q}_d = 2(d-1)$ and it is achieved with the maximally entangled state of two qudits and the following observables

$$A_x = U_x^\dagger F \Omega F^\dagger U_x, \qquad B_y = V_y F^\dagger \Omega F V_y^\dagger \tag{C.2.10}$$

where $\Omega = \mathrm{diag}[1, \omega, \ldots, \omega^{d-1}]$, F is the Fourier matrix given by

$$F = \frac{1}{\sqrt{d}} \sum_{i,j=0}^{d-1} \omega^{ij} \left|i\right\rangle\!\left\langle j\right|, \tag{C.2.11}$$

and, finally, U_x and V_y are unitary rotations defined as

$$U_x = \sum_{j=0}^{d-1} \omega^{j\theta_x} \left|j\right\rangle\left\langle j\right|, \qquad V_y = \sum_{j=0}^{d-1} \omega^{j\zeta_y} \left|j\right\rangle\left\langle j\right| \tag{C.2.12}$$

with $\theta_1 = 1/4$, $\theta_2 = 3/4$, and $\zeta_1 = 1/2$ and $\zeta_2 = 1$. The eigenprojectors of Alice's and Bob's measurements A_x and B_y are given by

$$\left|a\right\rangle_x = \frac{1}{\sqrt{d}} \sum_{k=0}^{d-1} \exp\left[i2\pi k(a-\theta_x)/d\right] \left|k\right\rangle \tag{C.2.13}$$

and

$$\left|b\right\rangle_y = \frac{1}{\sqrt{d}} \sum_{k=0}^{d-1} \exp\left[i2\pi k(-b+\zeta_y)/d\right] \left|k\right\rangle, \tag{C.2.14}$$

respectively, where $a, b \in \{0, \ldots, d-1\}$.

It is worth mentioning that by virtue of the Fourier transformation (C.2.7) the Bell expression (C.2.8) can be equivalently stated in the probability picture as

$$I_d := \sum_{k=0}^{\lfloor d/2 \rfloor - 1} \left(\alpha_k \mathbb{P}_k - \beta_k \mathbb{Q}_k\right), \tag{C.2.15}$$

where the expressions $\mathbb{P}_k$ and $\mathbb{Q}_k$ are defined as

$$\mathbb{P}_k = P(A_1 = B_1+k)+P(B_1 = A_2+k)+P(A_2 = B_2+k)+P(B_2 = A_1+k+1) \quad \text{(C.2.16)}$$

and

$$\mathbb{Q}_k = P(A_1 = B_1-k-1)+P(B_1 = A_2-k-1)+P(A_2 = B_2-k-1)+P(B_2 = A_1-k), \quad \text{(C.2.17)}$$

whereas the coefficients α_k and β_k are given by

$$\alpha_k = \frac{1}{2d}\left[g(k) + (-1)^d \tan\left(\frac{\pi}{4d}\right)\right], \qquad \beta_k = \frac{1}{2d}\left[g\left(k+1/2\right) - (-1)^d \tan\left(\frac{\pi}{4d}\right)\right], \quad \text{(C.2.18)}$$

with $g(k) = \cot[\pi(k+1/4)/d]$. Here, $P(A_x = B_y)$ is the probability that upon measuring A_x and B_y, Alice and Bob would obtain the same results, and it is explicitly given by

$$P(A_x = B_y) = \sum_{a=0}^{d-1} p(a, a|x, y). \quad \text{(C.2.19)}$$

Notice that by replacing the above α_k, β_k with $\alpha_k = \beta_k = 1 - 2k/(d-1)$ one obtains the CGLMP Bell expression [6].

The experimental violations of the multidimensional SATWAP Bell inequality $\tilde{I}_d \geq \mathcal{C}_d$ obtained in the experiment in chapter 3 are reported in Fig. 3.8 for different dimensionalities.

C.3 Steering inequalities for multidimensional bipartite maximally entangled states.

In the one-sided device-independent scenario both the measuring devices used by Alice and the quantum state shared between Alice and Bob are uncharacterised (or untrusted). The observable data in a so-called steering test are the probability distribution of Alice's measurement outcomes, conditional on her choice of measurement $p(a|x)$, and the states prepared, or 'steered' for Bob, given Alice's measurement and the shared state $\hat{\rho}_{a,x}$. In particular, assuming that the true state shared between Alice and Bob is $\hat{\rho}_d$, and that Alice's true measurements have POVM elements $\hat{M}_{a|x}$, then

$$\begin{aligned} p(a|x) &= \mathrm{Tr}[(\hat{M}_{a|x} \otimes \mathbb{I})\hat{\rho}_d], \\ \hat{\rho}_{a|x} &= \mathrm{Tr}_A[(\hat{M}_{a|x} \otimes \mathbb{I})\hat{\rho}_d]/p(a|x). \end{aligned}$$

This data is said to have a 'local-hidden-state' (LHS) model if there exists a hidden variable λ and collection of hidden states $\hat{\rho}_\lambda$, distributed according to $\varrho(\lambda)$, along with a response function $p(a|x,\lambda)$ that is able to reproduce it. In particular, if

$$p(a|x) = \int d\lambda \varrho(\lambda) p(a|x,\lambda) \tag{C.3.1}$$

$$\hat{\rho}_{a,x} = \frac{\int d\lambda \varrho(\lambda) p(a|x,\lambda)\hat{\rho}_\lambda}{\int d\lambda \varrho(\lambda) p(a|x,\lambda)} \tag{C.3.2}$$

Such a model provides a classical explanation of the steered states: all correlations between Alice's outcomes and Bob's states are mediated by the hidden variable λ, which controls Alice's measurements through the response function $p(a|x,\lambda)$, and Bob's state through $\hat{\rho}_\lambda$.

LHS models are not able to reproduce all the predictions of quantum theory. In particular, all LHS models satisfy linear constraints, known as EPR steering inequalities [7]. In this work we are interested in a particular class of EPR steering inequalities, specified by a choice of measurement operators for Bob, $\{\hat{M}_{b|y}\}_{b,y}$, of the form

$$\beta_d = \sum_{\substack{a=b \\ x=y}} p(a|x) \mathrm{Tr}\left[\hat{M}_{b|y}\hat{\rho}_{a|x}\right] \leq \beta_{lhs} \tag{C.3.3}$$

which can be violated by data $p(a|x)$, $\hat{\rho}_{a|x}$ arising in quantum theory. When $\{\hat{M}_{b|y}\}_{b,y}$ correspond to two mutually-unbiased-bases in dimension d, then $\beta_{lhs} = 1 + 1/\sqrt{d}$. Violations of EPR steering inequalities certify, in a one-sided device-independent manner, that the underlying state shared by Alice and Bob is entangled. Experimental values obtained for β_d, up to local dimension 15, are shown in Fig. 3.9.

C.4 Compressed sensing tomography

Quantum state tomography (QST) is a process which allows us to experimentally reconstruct the full density matrix of a quantum state. While QST represents an important diagnostic tool in the development of quantum technologies, its practical implementation becomes extremely challenging already for systems of intermediate scale. This is because the density matrix describing the joint state of n subsystems with local dimension d contains d^{2n} entries, which explodes exponentially with the system size.

In order to implement QST in near-term quantum technologies of increasing size, new techniques have been developed in recent years. Quantum compressed sensing is one of the most prominent ones.

Compressed sensing was originally developed for classical data analysis. The idea is to

exploit structures in the data collected in realistic situations to significantly reduce the number of parameters to be determined in a matrix reconstruction problem. Low-rank matrix recovery methods have inspired techniques for the reconstruction of low entropy quantum states, namely compressed sensing quantum state tomography (CSQST). These methods provide speed ups both in term of number of measurement required, and in the post-processing time required reconstruct the density matrix from the data [8–10], and have been implemented in systems of up to seven qubits [11]. In the experiment reported in chapter 3, the low entropy of the quantum states generated is also certified by the high quality of the quantum correlations measured with the protocols described above, which further justifies the use of CSQST.

The CSQST protocol requires projective measurements in the eigenbases of m operators randomly sampled from a chosen ortho-normal basis $\{\omega_i\}_{i=1}^{d^{2n}}$ of the operators space. Indicating with a vector $\boldsymbol{y}$ the measured statistics for all possible outcomes of these bases, and with $\mathcal{A}(\mathcal{X})$ the list of expectation values of the outcomes for a generic input $\mathcal{X}$, where $\mathcal{A}$ is a linear operator formalising the measurement process [10], the matrix reconstruction is performed by solving a the semi-definite program (SDP)

$$\begin{aligned} \tilde{\rho} = & \text{argmin } \mathrm{Tr}\mathcal{X} \\ & \text{s.t.} \quad \mathcal{X} \geq 0, \ \|\mathcal{A}(\mathcal{X}) - \boldsymbol{y}\|_2^2 \leq \epsilon, \end{aligned} \tag{C.4.1}$$

and finally renormalising $\rho = \tilde{\rho}/\mathrm{Tr}(\tilde{\rho})$. This is a convex trace norm minimisation constrained on semi-positive definite matrices, for which efficient convex optimization programs exist [12, 13]. Here ϵ is a control parameter related to the amount of noise in the data, which can be chosen according to the heuristics developed in Ref. [11]. In order to reconstruct a $N \times N$ density matrix of rank r, the measurements of $O(Nr\log(N^2))$ operators is sufficient, in contrast with $O(N^2)$ used for standard approaches [8–10]. In the low-rank case, this achieves approximately a quadratic advantage.

In the experiment reported in chapter 3, a total of 50, 122 and 228 operators were measured for dimensions 4, 8, and 12, respectively, following reported measurement procedures [11, 14]. The SDP for the density matrix reconstruction was performed using the SeDuMi solver [13] on a standard laptop, and took few seconds for $d = 4$, approximately five minutes for $d = 8$ and approximately one hour for $d = 12$. The obtained density matrices are shown in Fig. 3.10.

References

1. Sikora, J., Varvitsiotis, A. & Wei, Z. Minimum dimension of a Hilbert space needed to generate a quantum correlation. *Phys. Rev. Lett.* **117,** 060401 (2016).
2. Mermin, N. D. Simple unified form for the major no-hidden-variables theorems. *Phys. Rev. Lett.* **65,** 3373–3376 (1990).
3. Peres, A. Incompatible results of quantum measurements. *Phys. Lett. A* **151,** 107–108 (1990).
4. Aravind, P. Bell's theorem without inequalities and only two distant observers. *J. Genet. Couns.* **15,** 397–405 (2002).
5. Salavrakos, A., Augusiak, R., Tura, J., Wittek, P., Acín, A. & Pironio, S. Bell inequalities tailored to maximally entangled states. *Phys. Rev. Lett.* **119,** 040402 (2017).
6. Collins, D., Gisin, N., Linden, N., Massar, S & Popescu, S. Bell inequalities for arbitrarily high-dimensional systems. *Phys. Rev. Lett.* **88,** 040404 (2002).
7. Cavalcanti, D. & Skrzypczyk, P. Quantum steering: a review with focus on semidefinite programming. *Rep. Prog. Phys.* **80,** 024001 (2017).
8. Gross, D., Liu, Y.-K., Flammia, S. T., Becker, S. & Eisert, J. Quantum state tomography via compressed sensing. *Phys. Rev. Lett.* **105,** 150401 (2010).
9. Gross, D. Recovering low-rank matrices from few coefficients in any basis. *IEEE Trans. on Information Theory* **57,** 1548–1566 (2011).
10. Flammia, S. T., Gross, D., Liu, Y.-K. & Eisert, J. Quantum Tomography via Compressed Sensing: Error Bounds, Sample Complexity, and Efficient Estimators. *New J. Phys.* **14,** 095022 (2012).
11. Riofrío, C. A., Gross, D, Flammia, S. T., Monz, T, Nigg, D, Blatt, R & Eisert, J. Experimental quantum compressed sensing for a seven-qubit system. *Nat. Commun.* **8,** 15305 (2017).
12. Löfberg, J. YALMIP : A Toolbox for Modeling and Optimization in MATLAB. *2004 IEEE International Symposium on Computer Aided Control Systems Design (CACSD)* (2004).
13. Sturm, J. *Optim. Methods Softw.* **11-12,** 625–653 (1999).
14. Bolduc, E., Gariepy, G. & Leach, J. Direct measurement of large-scale quantum states via expectation values of non-Hermitian matrices. *Nat. Commun.* **7,** 10439 (2016).

D TEST MODELS FOR THE VALIDATION OF GAUSSIAN BOSON SAMPLING

In this appendix we provide a detailed description of the model used in chapter 4 for the validation of Gaussian boson sampling. For all models we adopted Bayesian approaches, which, as discussed in chapter 4, allow us to validate experimental data by comparing an ideal experimental implementation of GBS against a general test model for which output probabilities are computable. Here we describe how to compute such probabilities in some physically relevant models. Theorical contributions to the matierial reported here were provided by Levon Chakhmakhchyan.

D.1 Gaussian boson sampling with coherent states

We first consider a set of m input coherent states $\otimes_{i=1}^{m} |\alpha_i\rangle$, each of which has the following representation in the Fock basis,

$$|\alpha_i\rangle = e^{-\frac{|\alpha|^2}{2}} \sum_{n_i=0}^{\infty} \frac{\alpha^{n_i}}{\sqrt{n_i!}} |n_i\rangle \,. \tag{D.1.1}$$

A linear optical circuit U transforms a tensor product of coherent states $\otimes_{i=1}^{m} |\alpha_i\rangle$ into another tensor product of coherent states $\otimes_{i=1}^{m} |\beta_i\rangle$ with amplitudes

$$\beta_k = \sum_{j=1}^{M} U_{kj} \alpha_j. \tag{D.1.2}$$

In other words, coherent states remain separable while evolved through a linear optical circuit. Thus, the joint probability $p_{\text{cs}}(\mathbf{k})$ of detecting a pattern of n single photons $\mathbf{k} = \{k_1, \ldots, k_m\}$, with an m-mode coherent state $\otimes_{i=1}^{m} |\alpha_i\rangle$ at the input, admits a simple product form

$$p_{\text{cs}}(\mathbf{k}) = \prod_{i=1}^{m} \frac{e^{-|\beta_i|^2} |\beta_i|^{2k_i}}{k_i!}, \tag{D.1.3}$$

which is a product of Poisson distributions. Consequently, both sampling from it and computing its elements can be done efficiently on a classical computer. Although this problem is trivial from a computational point of view, boson sampling from coherent states comprises a physically relevant test model for validating our GBS experiments. In the model implemented for the experiment in chapter 4, the input coherent states amplitudes were considered to be uniform between all modes, as is physically plausible.

D.2 Gaussian boson sampling with thermal states

The next test model deals with sampling from linear-optically evolved thermal states. We consider an M-mode input thermal state $\otimes_{i=1}^{m} \rho_i^{\text{th}}$. In turn, each state ρ_i^{th} can be expressed as an incoherent mixture of Fock states,

$$\rho_i^{\text{th}} = (1 - \tau_i) \sum_{n_i=0}^{\infty} \tau_i^{n_i} |n_i\rangle \langle n_i| , \tag{D.2.1}$$

where $\tau_i = \langle n_i \rangle (\langle n_i \rangle + 1)$ and $\langle n_i \rangle$ is the average photon number of the state. The probability to detect a set $\mathbf{k} = \{k_1, \ldots, k_m\}$ of n single photons at the output of a linear-optical circuit U injected with an m-mode thermal state $\otimes_{i=1}^{m} \rho_i^{\text{th}}$ then reads

$$p_{\text{th}}(\mathbf{k}) = \frac{1}{k_1! \cdots k_m!} \frac{\text{Perm}\, A_{\mathbf{k}}}{\prod_{i=1}^{m} (1 + \langle n_i \rangle)}, \tag{D.2.2}$$

where

$$A = UDU^{\dagger}, \tag{D.2.3}$$

$$D = \text{diag}\{\tau_1, ..., \tau_m\} \tag{D.2.4}$$

and $A_{\mathbf{k}}$ is obtained from A by repeating k_i times its ith column and row.

Interestingly, although the photon-counting probability distribution for input thermal states is given in terms of matrix permanents (namely, of Hermitian positive semi-definite matrix permanents, since $\tau_i \geq 0$, $\forall i$) sampling from the probability distribution in Eq. (D.2.2) is classically tractable [1]. Even more, computing its elements can also be

done efficiently for a restricted set of linear-optical transformations U and input states ρ_i^{th} [2].

D.3 Gaussian boson sampling with distinguishable squeezed states

The treatment of GBS with distinguishable squeezed states is analogous to that of standard boson sampling with distinguishable single photons. Namely, since distinguishable squeezed states do not interfere with each other, the corresponding GBS experiment with k distinguishable squeezed states is equivalent to a set of k experiments where a single squeezed state $|\tilde{\xi}_j\rangle := |0,\dots,0,\xi_j,0,\dots,0\rangle$ evolves according to the m-mode linear-optical transformation U ($j=1,\dots,k$). In turn, the photon-counting statistics at its output is obtained by accumulating photon detection events from these experiments. To find the corresponding probability distribution, we first write down the Q-covariance matrix of the evolved states $|\tilde{\xi}_i\rangle$:

$$\tilde{\sigma}_{\text{Q}}^{(j)} = \frac{1}{2}\begin{bmatrix} U & 0 \\ 0 & U^* \end{bmatrix} \tilde{S}_j \tilde{S}_j^\dagger \begin{bmatrix} U^\dagger & 0 \\ 0 & U^{\text{T}} \end{bmatrix} + I_{2m}/2 \tag{D.3.1}$$

where

$$\tilde{S}_j = \begin{bmatrix} \oplus_{i=1}^m (1+\delta_{j,i}\cosh\xi_j) & \oplus_{i=1}^m \delta_{j,i}\sinh\xi_j \\ \oplus_{i=1}^m \delta_{j,i}\sinh\xi_j & \oplus_{i=1}^m (1+\delta_{j,i}\cosh\xi_j) \end{bmatrix}. \tag{D.3.2}$$

Next, in order to find the probability $p_{\text{dss}}(\mathbf{k})$ of detecting a pattern $\mathbf{k}=\{k_1,\dots,k_m\}$ of n single photons at the output of k distinguishable squeezed states distributed among m modes, one has to consider all the possible detection events upon the evolved states $|\tilde{\xi}_j\rangle$, $\forall j$, that yield the given pattern $\mathbf{k}$. Since the corresponding expression is rather bulky for arbitrary n, k and m, we present here only the expression relevant to the experiment discussed in chapter 4. That is, we assume that four distinguishable squeezed states are injected in the ath, bth, cth and dth mode of a 12-mode linear-optical circuit. The probability $p_{\text{dss}}^{(4)}(\mathbf{k})$ of detecting $n=4$ single photons in the qth, rth, sth and tth modes (no more than two photons per mode) at its output then reads

$$\begin{aligned} p_{\text{dss}}^{(4)}(\mathbf{k}) = & \frac{1}{\prod_{f\in\{a,b,c,d\}}\sqrt{\det\tilde{\sigma}_{\text{Q}}^{(f)}}}\Bigg(\sum_{\{i,j\}\in G}\Bigg[\frac{p_{l_1,l_2}^{(i)}p_{l_3,l_4}^{(j)}}{(1+\delta_{l_1,l_2})(1+\delta_{l_3,l_4})} + \frac{p_{l_1,l_3}^{(i)}p_{l_2,l_4}^{(j)}}{(1+\delta_{l_1,l_3})(1+\delta_{l_2,l_4})} \\ & + \frac{p_{l_1 l_4}^{(i)}p_{l_2,l_3}^{(j)}}{(1+\delta_{l_1,l_4})(1+\delta_{l_2,l_3})}\Bigg] + \sum_{i\in\{a,b,c,d\}}\left|\text{Haf}\,\tilde{C}_i'\right|^2\Bigg). \end{aligned} \tag{D.3.3}$$

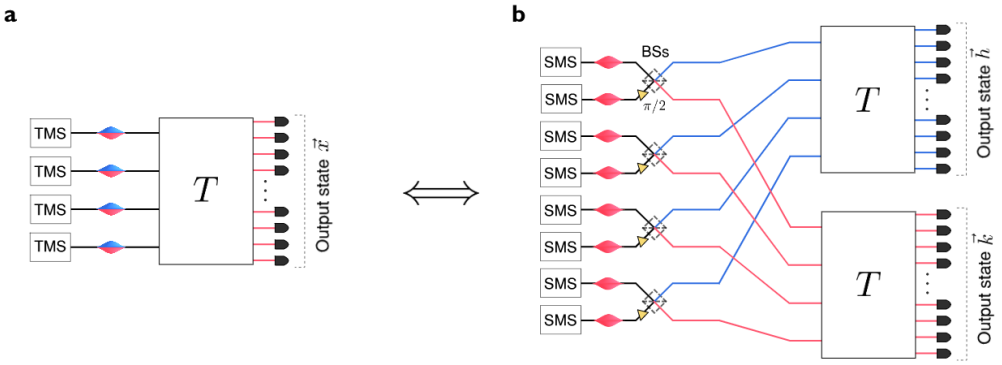


Figure D.1: Equivalence between a sampler with two-mode squeezed states at the input and a circuit with single-mode input squeezed states. (a) Photonic circuit in of a Gaussian boson sampler using TMS at the input. (b) Equivalent circuit with SMS at the input, where each TMS source is substituted with two SMS sources that are interfered into a beam-splitter after the bottom source has accumulated a $\pi/2$ phase.

Here, G is the set of all six distinct pairs of input mode numbers, that is

$$G = [\{a, b\}, \{a, c\}, \{a, d\}, \{b, c\}, \{b, d\}, \{c, d\}], \tag{D.3.4}$$

l is the list of mode numbers where photons were detected, $l = \{q, r, s, t\}$, $\tilde{C}'_i$ is obtained from the 12×12 matrix $\tilde{C}_i = U \cdot \mathrm{diag}[0, \ldots, \tanh \xi_i, \ldots, 0] \cdot U^{\mathrm{T}}$ by keeping its qth, rth, sth and tth rows and columns and $p_{ij}^{(j)} = |\tilde{C}_{ij}^{(j)}|^2$. If two photons were detected at a given mode, the corresponding mode number is repeated twice in the list l. For instance, if two photons were detected in the rth mode, the list l reads $l = \{q, r, r, t\}$. Note also that the first line in Eq. (D.3.3) corresponds to events where a pair of squeezed states produced a detection of two pairs of photons, while the second line corresponds to events where a single squeezed state produced a detection of four photons.

Although the computation of each element of the photon-counting probability distribution $p_{\mathrm{dss}}(\mathbf{k})$ is hard, sampling from it is classically tractable, analogous to the case of standard boson sampling with distinguishable single-photons (see, e.g., Ref. [3]).

D.4 Gaussian boson sampling with two-mode squeezed states

To treat the case where the input states are two-mode squeezed states instead of single-mode ones, we can make use of the analogy between the photonic circuits represented in Fig. D.1a and Fig. D.1b. Similar connections between sampling protocols using TMS and

SMS have already been highlighted in previous works, for example in Ref. [4] it was used to relate SBS and GBS. The idea is to note that m two-mode squeezers can be obtained from $2m$ single-mode squeezers combined pairwise into phase shifters and beam-splitters, as shown in Fig. D.1b. As photons in different modes of the two-mode squeezers would not interfere in the interferometer, we can separate all the m top output modes of the beam-splitters and the m bottom output modes, and send them into separate but equal interferometers T. Suppose that single-photon detection at the outputs of both interferometers, obtaining patterns $\mathbf{h}$ and $\mathbf{k}$ in the top and bottom m output modes, respectively (note that $\sum_i h_i = \sum_i k_i = n/2$ must hold, where n is the total number of photons). This would correspond to a detection patter $\mathbf{x} = \mathbf{k} + \mathbf{h}$ in the original TMS sampler. The output probabilities of the TMS sampler $p_{\text{tms}}(\mathbf{x})$ can then be calculated from the probabilities of the analogue SMS Gaussian boson sampler $p_{\text{tms}}(\mathbf{h}, \mathbf{k})$ as

$$p_{\text{tms}}(\mathbf{x}) = \sum_{\substack{\mathbf{h},\mathbf{k}: \\ \mathbf{x}=\mathbf{h}+\mathbf{k}, \\ \sum_i h_i = \sum_i k_i = n/2}} p_{\text{gbs}}(\mathbf{h}, \mathbf{k}), \tag{D.4.1}$$

where $p_{\text{gbs}}(\mathbf{h}, \mathbf{k})$ is the calculated probability of obtaining patterns $\mathbf{h}$ and $\mathbf{k}$ in the top and bottom m output modes of the scheme in Fig. D.1b. The probability $p_{\text{gbs}}(\mathbf{h}, \mathbf{k})$ is obtained using the total transfer matrix $T_{\text{gbs}} = (T \oplus T) \cdot T_{\text{bs}} \cdot T_{\text{ps}}$ which includes the matrices associated to the array of phase shifters $T_{\text{ps}} = U_{\text{ps}}^{\oplus m}$ and beam-splitters $T_{\text{bs}} = U_{\text{bs}}^{\oplus m}$, where

$$U_{\text{ps}} = \begin{bmatrix} 1 & 0 \\ 0 & i \end{bmatrix}, \qquad U_{\text{bs}} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} / \sqrt{2}. \tag{D.4.2}$$

We remark that, although sampling from TMS input states deviates from an ideal implementation of GBS, it still represents a task which is hard classically, as it can be mapped into an analogous GBS problem. Therefore, rather then a validation on the computational complexity of the sampler, the test performed should here be interpreted as supporting evidence for the correct experimental implementation of the protocol.

References

1. Rahimi-Keshari, S., Lund, A. P. & Ralph, T. C. What can quantum optics say about computational complexity theory? *Phys. Rev. Lett.* **114,** 060501 (2015).
2. Chakhmakhchyan, L., Cerf, N. J. & Garcia-Patron, R. A quantum-inspired algorithm for estimating the permanent of positive semidefinite matrices. *Phys. Rev. A* **96,** 022329 (2017).
3. Aaronson, S. & Arkhipov, A. The computational complexity of linear optics. *Quant. Inf. Comp.* **13,** 1383 (2014).
4. Kruse, R., Hamilton, C. S., Sansoni, L., Barkhofen, S., Silberhorn, C. & Jex, I. A detailed study of Gaussian Boson Sampling. *arXiv:1801.07488* (2018).